



Etude, analyse et modélisation physique de la production de la parole avec applications aux troubles liés à une surdité profonde

Louis Delebecque

► To cite this version:

Louis Delebecque. Etude, analyse et modélisation physique de la production de la parole avec applications aux troubles liés à une surdité profonde. Traitement du signal et de l'image [eess.SP]. Université Grenoble Alpes, 2015. Français. NNT : 2015GREAT118 . tel-01269639

HAL Id: tel-01269639

<https://theses.hal.science/tel-01269639>

Submitted on 5 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse

pour obtenir le grade de

Docteur de l'université Grenoble Alpes

Spécialité : **Signal Image Parole Télécoms**

Arrêté ministériel : 7 août 2006

Présentée par

Louis DELEBECQUE

Thèse dirigée par **Xavier PELORSON** et
codirigée par **Denis BEAUTEMPS**

préparée au sein du **laboratoire GIPSA-Lab**
dans l'école doctorale **Électronique, Électrotechnique,
Automatique, Traitement du Signal (EEATS)**

**Étude, analyse et modélisation physique de la
production de la parole avec applications aux
troubles liés à une surdité profonde**

Thèse soutenue publiquement le **21 Septembre 2015**,
devant le jury composé de :

Christophe d'ALESSANDRO

DR CNRS, LIMSI, Orsay, Rapporteur

Avraham HIRSCHBERG

Prof. émérite, Technical University of Eindhoven, Rapporteur

Joël GILBERT

DR CNRS, LAUM, Le Mans, Examineur

Christophe VERGEZ

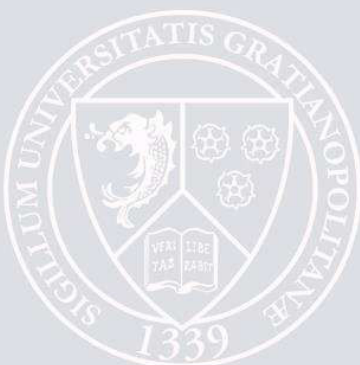
DR CNRS, LMA, Marseille, Examineur (Président du Jury)

Xavier PELORSON

DR CNRS, GIPSA-lab, Grenoble, Directeur de thèse

Denis BEAUTEMPS

CR CNRS, GIPSA-lab, Grenoble, Codirecteur de thèse



UNIVERSITÉ GRENOBLE APLES
ÉCOLE DOCTORALE EEATS
Électronique, Électrotechnique, Automatique, Traitement du Signal

THÈSE

pour obtenir le titre de

docteur en sciences

de l'Université Grenoble Alpes

Mention : SIGNAL IMAGE PAROLE TÉLÉCOMS

Présentée par

Louis DELEBECQUE

**Étude, analyse et modélisation physique de la production de la
parole avec applications aux troubles liés à une surdité profonde**

Thèse dirigée par Xavier PELORSON et Denis BEAUTEMPS

préparée au laboratoire GIPSA-Lab

soutenue le 21 Septembre 2015

Jury :

<i>Rapporteurs :</i>	Christophe d'ALESSANDRO	-	LIMSI, Orsay
	Avraham HIRSCHBERG	-	Technical University of Eindhoven
<i>Directeur :</i>	Xavier PELORSON	-	GIPSA-lab, Grenoble
<i>Codirecteur :</i>	Denis BEAUTEMPS	-	GIPSA-lab, Grenoble
<i>Examineur :</i>	Joël GILBERT	-	LAUM, Le Mans
	Christophe VERGEZ	-	LMA, Marseille (Président du Jury)

Remerciements

Je souhaite remercier vivement Christophe d'Alessandro et Mico Hirschberg pour l'intérêt qu'ils ont porté à ce travail en acceptant de le rapporter. J'aimerais également remercier Christophe Vergez et Joël Gilbert qui ont accepté de l'examiner.

Je remercie Xavier Pelorson et Denis Beautemps pour avoir initié ces travaux et pour m'avoir fait confiance tout au long de ce travail de thèse. La partie expérimentale de ce travail a bénéficié de l'appui logistique de Xavier Laval et Christophe Savariaux tout deux responsables de plate-formes expérimentales au département Parole et Cognition du Gipsa-lab.

Je souhaite également remercier le professeur Sébastien Schmerber, médecin ORL ainsi que Denise Martin-Bonniot, orthophoniste au CHU Grenoble-Alpes, de m'avoir mis en contact avec l'unique locuteur sourd ayant participé à ces travaux de thèse. Je remercie également ce dernier pour sa bonne volonté et son grand intérêt pour ces recherches.

Merci également aux enseignements-chercheurs du département Génie Civil de l'IUT1 de Grenoble et de l'université du Maine, avec lesquels j'ai eu l'occasion de travailler pour la préparation d'enseignements. Je pense en particulier à Claire Brutel-Vuilmet, Olivier Dazel, Guillaume Penelet, Bruno Brouard et Christophe Ayrault.

Durant mon séjour au département Parole et Cognition du Gipsa-Lab, j'ai pu bénéficier des conseils et indications de Fabrice Silva, Maëva Garnier, Coriandre Vilain, Nathalie Henrich, Pierre Badin et Thomas Hueber. Merci aussi aux doctorants et post-doctorants du département d'avoir agrémenter ma vie quotidienne au laboratoire, je pense en particulier à mes collègues du bureau « TAD » : Nico, Noël et Paul.

Ces trois années passées à Grenoble, n'aurait pas été aussi agréables sans mes premiers hôtes grenoblois : Toby, Stan et Illya, sans les colocataires du « manoir » de Corenc : Tatou, Emma, Cécile, Brunelle, Rémi, Maïté, Colin, Lynn, Séb, Yordie, Nicolas et Carole et sans les répétitions hebdomadaires avec les Georges du Gondwana composés de Riton, Colin, Adrien, Clémich et Féloch.

Enfin, je remercie ma famille et Balbine pour son soutien inconditionnel pendant ces années de thèse.

Table des matières

Listes des sons	vii
Introduction	1
I Partie préliminaire	5
1 Notions sur la production de sons voisés	7
1.1 Aspects physiologiques	8
1.2 Aspects physiques : Pressions de seuil des oscillations	12
1.3 Aspects phonétiques	15
1.4 Éléments sur la production de parole chez les sourds	18
1.5 Bilan	21
2 Modèles théoriques et implémentation numérique	23
2.1 Introduction	23
2.2 Modèles aérodynamiques	24
2.3 Modèles mécaniques	31
2.4 Modèles acoustiques	35
2.5 Implémentation numérique	49
2.6 Bilan	51
II Sauts de fréquence fondamentale et transitions de mécanisme laryngé	53
3 Bibliographie et observations <i>in vivo</i>	57
3.1 Introduction	57

3.2	Éléments bibliographiques	58
3.3	Observation <i>in vivo</i> de saut de fréquence fondamentale	62
4	Validation expérimentale par simulation numérique	69
4.1	Introduction	69
4.2	Dispositif expérimental	70
4.3	Résultats expérimentaux	73
4.4	Simulation numérique	77
4.5	Discussion et conclusion	87
5	Étude des stratégies motrices par simulation numérique	93
5.1	Introduction	93
5.2	Choix des paramètres de simulation	94
5.3	Méthode d'analyse des simulations	95
5.4	Résultats	97
5.5	Discussion et conclusion	105
III	Séquences voyelle - plosive non voisée - voyelle	109
6	Interaction aérodynamique entre les cordes et le conduit vocal	113
6.1	Introduction	114
6.2	Measurements on a human speaker	115
6.3	Air flow model for the lip constriction	118
6.4	Aerodynamic model of the supraglottal cavity expansion	123
6.5	Numerical simulations of an /apa/ sequence	128
6.6	Conclusion	130
7	Influence du couplage acoustique sur le VOT	133

7.1	Introduction	133
7.2	Mesures sur maquette de l'appareil phonatoire	136
7.3	Simulations numériques	138
7.4	Discussion et conclusion	141
Conclusion		145
A Couches limites et séparation de l'écoulement		151
B Calibration de la station EVA		155
C Mesures complémentaires de sauts de fréquence fondamentale		165
D Résonateurs acoustiques		167
E Cartographies des régimes d'oscillation		171
Bibliographie		177

Listes des sons

Note Liminaire : Les travaux présentés dans ce manuscrit sont illustrés par des fichiers audio. L'accès à ces fichiers, au format mp3, nécessite une connection internet. Les liens vers les fichiers audio sont symbolisés par un encadré. Les lecteurs dont le navigateur ne permet pas de lire directement les fichiers mp3 ont la possibilité de télécharger les fichiers audio en utilisant l'adresse URL donnée par chaque lien afin de les lire avec un lecteur multimédia indépendant.

- [son 1](#) Pression acoustique p_a mesurée lors de la production de la voyelle /u/ produite avec une forte intensité par un locuteur sourd, mesure présentée en figure 3.2
- [son 2](#) Pression acoustique p_{ac} mesurée lors de la variation de la pression d'eau contenue dans les cordes vocales artificielles, mesure présentée en figure 4.9
- [son 3](#) Pression acoustique p_{ac} mesurée lors de la variation de la pression en amont de la maquette de cordes vocales, mesure présentée en figure 4.5
- [son 4](#) Dérivée temporelle du débit glottique dU_g/dt simulé lors de la variation de la raideur des cordes vocales, sans couplage acoustique, simulation présentée en figures 4.9 et 4.12
- [son 5](#) Pression acoustique supra-glottique p_{av} simulée lors de la variation de la raideur des cordes vocales en prenant en compte le couplage acoustique, simulation présentée en figure 4.12
- [son 6](#) Dérivée temporelle du débit glottique dU_g/dt simulé lors de la variation de la pression sous-glottique, sans couplage acoustique, simulation présentée en figures 4.13 et 4.14.
- [son 7](#) Pression acoustique supra-glottique p_{av} simulée lors de la variation de la pression sous-glottique en prenant en compte le couplage acoustique, simulation présentée en figure 4.14.
- [son 8](#) Exemple de pression acoustique p_M rayonnée aux lèvres, représentatif du régime 1, pour une production de la voyelle /a/, régime définit en section 5.3.2.
- [son 9](#) Exemple de pression acoustique p_M rayonnée aux lèvres, représentatif du régime 2, pour une production de la voyelle /a/, régime définit en section 5.3.2.
- [son 10](#) Exemple de pression acoustique p_M rayonnée aux lèvres, représentatif du régime C, pour une production de la voyelle /a/, régime définit en section 5.3.2.
- [son 11](#) Pression acoustique P_a mesurée lors de la production de la séquence /apa/ par un locuteur normo-entendant, mesure présentée en figure 6.2.
- [son 12](#) Pression acoustique p_{av} simulée afin de reproduire la mesure de la séquence voyelle-plosive-voyelle réalisée sur la maquette de cordes vocales avec un tube en plexiglas de 10 cm, simulation présentée en figure 7.5
- [son 13](#) Pression acoustique p_{ac} mesurée lors de la production de la séquence voyelle-plosive-voyelle sur la maquette de cordes vocales avec un tube en plexiglas de 10 cm, mesure présentée en figures 7.3 et 7.5.

Introduction

Contexte

Physique de la production de sons voisés

La production des sons qui constituent le langage parlé résulte de multiples interactions fluide-structure complexes dans l'ensemble des voies aériennes supérieures. En particulier, la source voisée produite par l'auto-oscillation des cordes vocales résulte de l'interaction entre l'écoulement d'air provenant des poumons, le mouvement des tissus élastiques constituant les cordes vocales et les ondes acoustiques se propageant à travers la trachée et le conduit vocal. Ainsi, le comportement de la source de sons voisés est contrôlé par les muscles du larynx, par les mouvements du diaphragme, et par la géométrie des voies respiratoires supérieures qui influent sur la résistance au passage de l'air et sur les résonances acoustiques. Le contrôle de la source voisée est donc déterminé par l'ensemble des organes intervenant dans la production de la parole.

Production de la parole : contrôle et perception auditive

La production du langage parlé nécessite la réalisation de mouvements et d'ajustements précis des différents organes intervenant dans la production de la parole. La capacité d'effectuer et d'ajuster ces mouvements constitue le contrôle moteur. Selon Perkell [PML⁺97], lors de la production de parole, les objectifs d'un locuteur sont auditifs et le contrôle moteur permet de viser une cible acoustique associée à un son élémentaire du langage. Le contrôle moteur repose d'une part sur la rétroaction de la boucle de perception-production appelée *feedback* et d'autre part sur les modèles internes de production. Les modèles internes sont équivalents à un dictionnaire répertoriant l'ensemble des relations entre les commandes motrices articulatoires et les sons élémentaires du langage. La construction des modèles internes est alimentée durant la période d'acquisition du langage, dans l'enfance, par les *feedbacks* auditif, visuel et proprioceptif¹. Chez l'enfant, l'acquisition du langage passe par une phase d'imitation pendant laquelle il cherche à reproduire les sons qu'il entend et les gestes articulatoires qu'il voit. La modalité auditive est alors la plus importante dans la construction des modèles internes. Une fois que les modèles internes sont fonctionnels et robustes, le locuteur est capable de produire une cible acoustique précisément sans avoir besoin de l'intervention des différents *feedbacks*. La rétro-action a alors uniquement pour fonction d'adapter la production en cas de perturbation extérieure ponctuelle (par exemple celle d'ajuster le niveau sonore et le timbre de la voix lorsque l'environnement est bruyant) ou bien d'ajuster progressivement les modèles internes en cas d'une modification physiologique (par exemple la mue de la voix chez l'adolescent).

Production de la parole et surdité

1. Le terme proprioception désigne les sensations qu'a un individu des positions et mouvements de son propre corps. Dans le cadre de la production de la parole, la proprioception inclut la sensation tactile (de la langue sur le palais par exemple) et la sensation de la tension des muscles articulateurs.

Pour un locuteur devenu sourd ou malentendant après l'acquisition du langage (surdité post-linguistique), la disparition du *feedback* auditif provoque une dégradation progressive des modèles internes et donc de la précision du contrôle moteur de la production de parole. La dégradation des modèles internes est responsable de l'apparition de troubles liés à un défaut de contrôle. Ces troubles dus à une mauvaise gestion du souffle, de l'activité des muscles du larynx ou des ajustements articulatoires (et donc des résonances acoustiques), se manifestent au niveau phonologique (omission, ajout ou substitution de phonèmes) ou bien au niveau prosodique (rythme, hauteur et intensité). L'importance des troubles de la production est donc fortement liée au niveau d'audition et à l'âge du locuteur lors de l'apparition des pertes auditives.

Dans le cas des personnes devenues sourdes avant l'acquisition du langage (surdité pré-linguistique), la parole n'est pas un moyen de communication instinctif, d'autres modalités telles que la langue des signes sont privilégiées. Une éducation à la communication parlée leur est tout de même inculquée durant l'enfance (généralement) dans le but d'une meilleure insertion dans la société, il s'agit de l'oralisation. Les jeunes sourds construisent alors leurs modèles internes à partir des *feedbacks* visuel et surtout proprioceptif avec une aide orthophonique. Leur production de parole est donc régie par un contrôle moteur ne bénéficiant pas de *feedback* auditif et basé sur des modèles internes qui diffèrent de ceux des personnes normo-entendantes. Là encore, l'importance des troubles de la communication parlée dépend du degré de pertes auditives.

Par ailleurs, une proportion non négligeable des personnes atteintes de surdité pré ou post-linguistique d'âge et de degré de pertes auditives différents ont été implantées² ces deux dernières décennies. Dans la mesure où les modèles internes d'une telle population ont évolué ou se sont construits sans *feedback* auditif, où la qualité du son perçu diffère de celle obtenue par l'audition naturelle, le contrôle moteur de cette population, qui évolue en fonction du recouvrement de l'audition, peut également conduire à des troubles de la phonation.

Dans tous ces cas de surdité, le contrôle moteur altéré est la cause de troubles de la production qui peuvent nuire à la communication parlée et induire des difficultés d'ordre social.

Les études réalisées dans cette thèse ne concernent pas les origines d'un défaut de contrôle moteur mais ses effets. Ces troubles de la production peuvent notamment se manifester au niveau du contrôle de la source voisée de la parole, or la maîtrise de l'auto-oscillation des cordes vocales est essentielle à la qualité de la communication parlée. Par exemple, l'évolution de la fréquence fondamentale de la voix parlée a une fonction sémantique³, l'intonation déclarative en est un bon exemple. De plus, la production de nombreux phonèmes nécessite une coordination précise entre l'apparition ou l'arrêt du voisement et la production d'un son au niveau des voies aériennes supérieures.

2. L'implantation cochléaire est l'implantation chirurgicale dans l'oreille interne d'un dispositif qui envoie des signaux électriques au nerf auditif en fonction du signal capté par un microphone. Un implant cochléaire permet à un sujet sourd profond de percevoir un signal auditif.

3. Sémantique : qui donne du sens.

Objectifs

Les travaux réalisés dans cette thèse s'inscrivent dans un contexte de modélisation physique de la production de parole. L'objectif est de contribuer à la compréhension des phénomènes physiques gouvernant la production des sons voisés. À plus long terme, ces travaux s'inscrivent dans le développement d'un outil numérique de synthèse de parole par modèle physique.

La phonation d'une personne sourde offre alors un cadre d'étude intéressant. En effet, alors que l'appareil phonatoire ne présente *a priori* aucune anomalie, l'absence de *feedback* auditif permet de mettre en avant des comportements physiques de l'instrument vocal qui ne sont pas observés lorsque celui-ci est parfaitement maîtrisé. L'observation de troubles liés à un défaut de contrôle de la source voisée, dans un but d'amélioration des modèles physiques disponibles, constitue donc l'originalité de l'approche de ce travail.

Parallèlement, une bonne compréhension physique des phénomènes qui régissent la production de parole éclaire sur les différentes stratégies motrices, adaptées ou non, employées par la personne sourde, notamment lorsque des interactions fluide-structure complexes sont impliquées. La modélisation peut alors permettre de déterminer l'origine physique des troubles observés et ainsi d'élaborer des préconisations pour aider à les corriger.

Cette démarche vise également à établir des liens entre les connaissances physiques sur la production de la parole et les connaissances phonétiques pour deux cas concrets de production.

La première étude se concentre sur les sauts de fréquence fondamentale lors de la production d'une voyelle tenue. La deuxième étude porte sur la production de consonnes plosives, en particulier sur les effets de la réalisation d'une occlusion du conduit vocal sur l'arrêt puis l'apparition de l'oscillation des cordes vocales.

Méthode

L'approche globale adoptée dans cette thèse consiste, dans un premier temps, à observer les phénomènes étudiés au moyen de mesures *in vivo*. La capacité des modèles à reproduire les phénomènes observés est ensuite évaluée en utilisant des mesures réalisées sur une maquette simplifiée de l'appareil phonatoire. L'intérêt d'utiliser ce dispositif incluant une maquette auto-oscillante des cordes vocales, est que ses paramètres sont davantage contrôlables, observables et peuvent plus facilement être associés aux grandeurs intervenant dans le modèle physique. Enfin, les modèles théoriques sont utilisés pour réaliser des simulations numériques temporelles afin d'étudier l'effet des interactions physiques sur la source laryngée et d'expliquer les différentes stratégies motrices possibles pour la production de sauts de fréquence fondamentale et de consonnes plosives.

Le manuscrit est structuré en trois parties. La première partie permet d'une part d'introduire des notions relatives à la production de sons voisés et aux troubles de la production chez les locuteurs sourds et malentendants (Chapitre 1) et d'autre part de présenter les modèles théoriques qui alimentent les simulations numériques (Chapitre 2).

La deuxième partie traite des sauts de fréquence fondamentale observés lors d'une transition involontaire de mécanisme laryngé. Les connaissances bibliographiques sur ce sujet sont résumées puis le phénomène est observé chez un sujet sourd profond post-linguistique

implanté depuis peu (Chapitre 3). Ce phénomène de sauts de fréquence est reproduit sur une maquette auto-oscillante de l'appareil phonatoire. Ces mesures sont utilisées pour valider les simulations numériques basées sur les modèles théoriques proposés (Chapitre 4). Des simulations numériques sont ensuite utilisées pour imiter les productions d'un locuteur humain afin de déterminer les différentes stratégies motrices responsables des sauts de fréquence fondamentale (Chapitre 5).

La troisième partie porte sur la modélisation de la production de séquences voyelle - plosive bilabiale non voisée (/p/) - voyelle. Dans un premier temps, un modèle aérodynamique est proposé, validé puis utilisé pour étudier les effets de la formation d'une occlusion du conduit vocal sur l'arrêt puis la reprise du voisement (Chapitre 6). Dans un second temps, l'influence des interactions acoustiques entre le résonateur situé en aval du larynx et les cordes vocales sur le délai d'établissement du voisement après le relâchement de l'occlusion est déterminé (Chapitre 7).

Première partie

Partie préliminaire

Notions sur la production de sons voisés

Sommaire

1.1 Aspects physiologiques	8
1.1.1 Appareil phonatoire humain	8
1.1.2 Structure des cordes vocales	8
1.1.3 Muscles du larynx	9
1.1.4 Mécanismes laryngés	10
1.2 Aspects physiques : Pressions de seuil des oscillations	12
1.2.1 Définition	12
1.2.2 Hystérésis entre l'apparition et l'extinction des oscillations	12
1.2.3 Comportement de la pression P_{on}	13
1.3 Aspects phonétiques	15
1.3.1 Voyelles	15
1.3.2 Fréquence fondamentale intrinsèque des voyelles	16
1.3.3 Consonnes	16
1.3.4 Voice onset Time	17
1.4 Éléments sur la production de parole chez les sourds	18
1.4.1 Fréquence fondamentale	19
1.4.2 Nasalisation	19
1.4.3 Voyelles	20
1.4.4 Consonnes	20
1.5 Bilan	21

Ce premier chapitre permet de définir le cadre de cette thèse.

La production de sons voisés repose sur le phénomène d'auto-oscillation des cordes vocales. Les interactions fluide-structure au niveau des cordes vocales génèrent et entretiennent un mouvement périodique à partir de la source d'énergie continue que représente l'écoulement d'air provenant des poumons. Les notions relatives à la production de sons de la parole faisant intervenir ce phénomène d'oscillation auto-entretenu des cordes vocales sont introduites. Les aspects physiologiques, physiques et phonétiques de la production sont ici brièvement exposés. Les principaux troubles de la production de parole observés chez les locuteurs atteints de surdité partielle ou profonde sont également présentés.

1.1 Aspects physiologiques

1.1.1 Appareil phonatoire humain

La figure 1.1.a illustre la localisation des organes du corps humain intervenant dans la production de parole. Ces organes se répartissent sur trois niveaux ayant chacun une fonction dans la production.

Sous-glottique

Les poumons et la trachée constituent la soufflerie qui fournit la source d'énergie nécessaire à l'ensemble de l'appareil phonatoire. L'effort pulmonaire est régi par les mouvements du diaphragme. La parole est essentiellement produite lors de l'expiration de l'air, qui est à l'origine de la formation d'une surpression en dessous du larynx, appelée pression sous-glottique.

Glottique

Le larynx, situé au carrefour des voies respiratoire et œsophagienne, est la partie de l'appareil phonatoire qui abrite des replis de membranes muqueuses appelées cordes vocales (ou plis vocaux). L'espace compris entre les deux cordes vocales est appelée la glotte. Les sons voisés sont produits par la vibration des cordes vocales, qui modulent le débit d'air glottique.

Supra-glottique

Les voies aériennes supérieures se composent du conduit vocal et de la cavité nasale. L'ensemble des éléments du conduit vocal (langue, palais, lèvres, dents, épiglote) sont impliqués dans l'articulation, ainsi que le voile du palet qui permet lorsqu'il est abaissé de faire communiquer le volume du conduit vocal avec celui de la cavité nasale. L'articulation est la modulation du signal de la source voisée par les résonances acoustiques des cavités supra-glottiques.

L'interaction fluide-structure entre l'écoulement d'air provenant des poumons et l'élasticité des cordes vocales permet donc de produire des sons voisés. Ce voisement se traduit par l'apparition et l'entretien d'un mouvement de vibration périodique des cordes vocales. La source de débit acoustique ainsi créée par la modulation de l'écoulement glottique a une fréquence fondamentale, noté F_0 , imposée à la fréquence de vibration des cordes vocales.

1.1.2 Structure des cordes vocales

Les cordes vocales possèdent une structure constituée de plusieurs couches de composition différente. La partie la plus profonde est composée de fibres musculaires appartenant aux muscles thyro-aryténoïdiens (également appelés muscles vocaux). La couche moyenne, appelée *lamina propria*, est un tissu conjonctif lâche en partie constitué de fibres dont la densité augmente avec la profondeur. La couche superficielle est une muqueuse appelée l'épithélium.

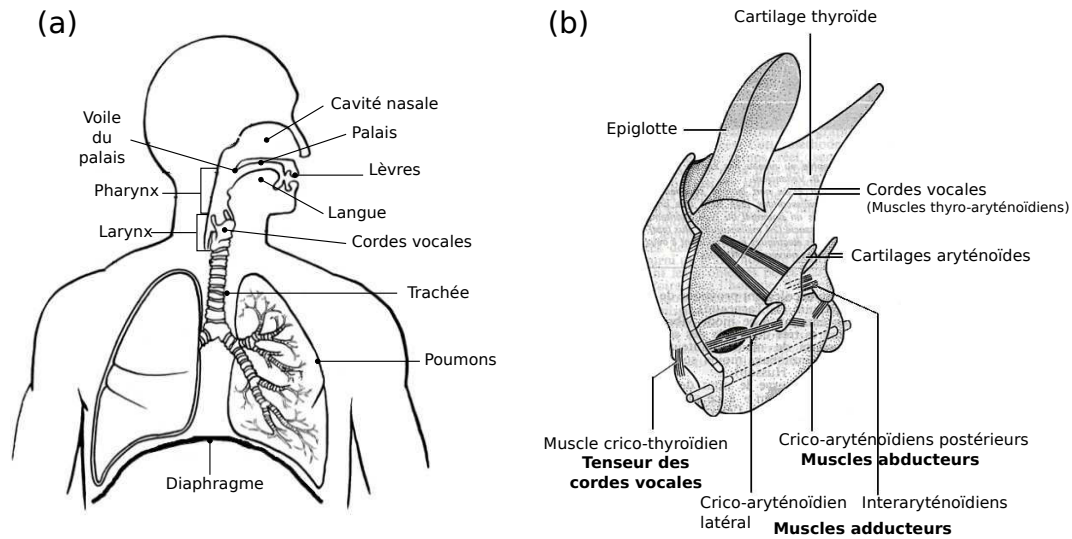


FIGURE 1.1 – (a) : Coupe de l'appareil phonatoire humain, schéma adapté de [Ken14]. (b) : Schéma de la mécanique du larynx (la trachée et le pharynx ne sont pas représentés), schéma adapté de [Lul53].

1.1.3 Muscles du larynx

Les muscles qui connectent les différents cartilages du larynx permettent de contrôler la tension et l'écartement des cordes vocales. La figure 1.1.b présente les cartilages du larynx et les muscles intervenant dans la production de parole.

Muscles tenseurs des cordes vocales

La contraction des muscles thyro-aryténoïdiens rigidifie les cordes vocales ce qui entraîne une augmentation de la fréquence fondamentale de la phonation. L'action des muscle crico-thyroïdiens provoque le basculement du cartilage thyroïde (pomme d'Adam) vers l'avant. Ce mouvement provoque une augmentation de la longueur et de la tension longitudinale des cordes vocales. Tout comme la contraction des muscles thyro-aryténoïdiens, cette stratégie est utilisée pour augmenter la fréquence fondamentale de la phonation.

Muscles adducteurs et abducteurs

Les cordes vocales sont écartées (abduction) lors de la respiration, afin de laisser passer l'air, et rapprochées lors de la phonation (adduction). Le niveau d'adduction des cordes vocales a un impact direct sur le timbre de la production vocale. Le geste d'adduction des cordes vocales est assuré par l'action des muscles inter-aryténoïdiens, dont la contraction a pour effet de rapprocher les deux cartilages aryténoïdes, et par la contraction des muscles crico-aryténoïdiens latéraux, qui fait pivoter les aryténoïdes. L'abduction des cordes vocales est réalisée par le biais des muscles crico-aryténoïdiens postérieurs dont la contraction provoque un basculement des aryténoïdes, qui a pour effet d'écarter les cordes vocales l'une de l'autre.

1.1.4 Mécanismes laryngés

Description des principaux mécanismes laryngés

Les productions vocales humaines peuvent être classées selon plusieurs registres qui diffèrent par le motif de vibration des cordes vocales lors de la phonation, par la gamme de fréquence fondamentale qui en résulte et par la manière dont ils sont perçus. Dans ce document, ces registres vocaux seront désignés par le terme de mécanisme laryngé proposé par Roubeau [Rou93].

Les productions vocales non-pathologiques comptent au total quatre mécanismes laryngés. Les caractéristiques des mécanismes I et II, qui sont les plus fréquemment rencontrés, sont décrites ci-dessous à partir des observations ou des informations rapportées dans les études suivantes [dBT59, Ven67, Rou93, Hen01].

Mécanisme I (MI)

Les cordes vocales sont épaisses et vibrent sur toute leur longueur, l'amplitude de la vibration est importante et la fermeture glottique est complète. La durée de la phase fermée est importante, environ égale à celle de la phase ouverte. Le mécanisme I est le mécanisme couramment employé en voix parlée. D'un point de vue acoustique, il se caractérise également par un timbre riche en raison de l'effet de la collision prononcée entre les cordes vocales, sur la source laryngée. Le mécanisme I correspond au registre de la voix modale, également appelée voix de poitrine.

Mécanisme II (MII)

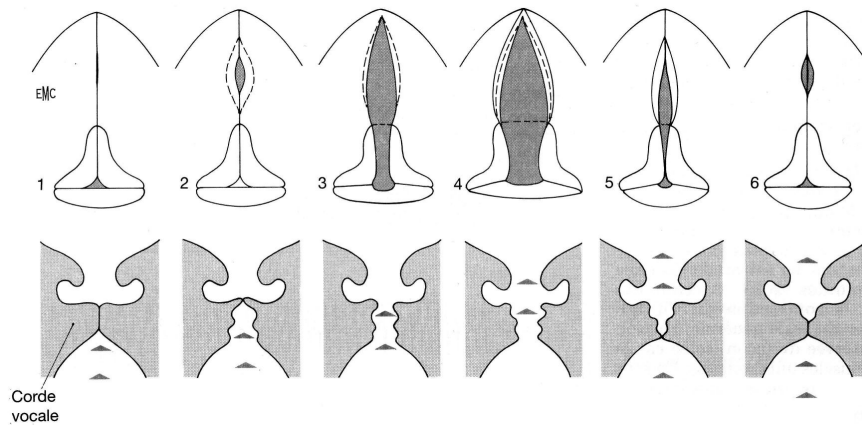
Les cordes vocales sont plus fines car la partie interne constituée du muscle thyroïdien-aryténoïdien ne participe quasiment pas à la vibration. La vibration s'effectue uniquement sur 2/3 de la longueur des cordes vocales. L'amplitude de la vibration est plus faible et la fermeture glottique est incomplète. La durée de la phase ouverte est plus importante que celle de la phase fermée. Le timbre produit par la phonation en mécanisme II est relativement pauvre. Le mécanisme II correspond au registre vocal appelé *falsette* ou encore voix de tête.

Les différences entre les motifs de vibration des cordes vocales pour les mécanismes I et II sont illustrées par la figure 1.2.

Les mécanismes laryngés I et II se différencient essentiellement par la façon dont se produit la collision entre les cordes vocales opposées. En MI, les forces mises en jeu lors de la collision sont plus importantes. Ce phénomène a deux conséquences : la surface de contact est plus importante en MI et le quotient ouvert Q_o , défini comme le rapport entre la durée de la phase ouverte et la période des oscillations, est plus élevé en MII. En voix chantée, les valeurs de Q_o sont comprises entre 30 et 80 % en MI et entre 50 et 95 % en MII [Hen01].

Au vu de ces observations, l'électroglottographie se présente comme une technique adaptée à l'identification des mécanismes laryngés.

Mécanisme I



Mécanisme II

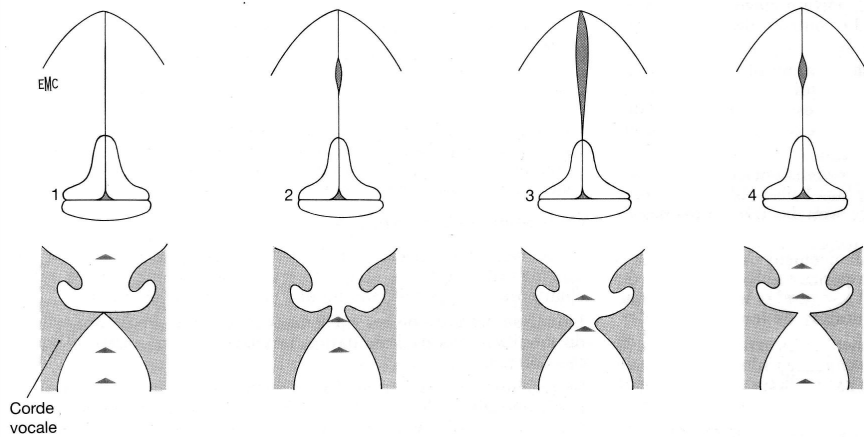


FIGURE 1.2 – Représentation des cycles de vibration glottique en mécanisme I et II, dans les plan horizontal (vue du dessus) et dans le plan coronal, pendant une période fondamentale, d'après [Ven67] et adapté de [Hen01].

Électroglottographie

L'électroglottographie est une méthode indirecte pour l'étude de la vibration glottique. Par l'intermédiaire de deux électrodes placées sur le cou de part et d'autre de la glotte du sujet, l'électroglottographe (*EGG*) [Fab57] réalise une mesure instantanée de la conductivité du larynx selon l'axe transversal. La variation de la conductivité peut être assimilée à la variation de la surface de contact entre les cordes vocales opposées. La dérivée temporelle du signal *EGG* est appelée *DEGG*. Le signal *DEGG* est utilisé pour déterminer les instants d'ouverture t_o et de fermeture t_f des cordes vocales qui correspondent respectivement aux maximums des vitesses d'augmentation et de réduction du contact entre les cordes vocales. Ces informations permettent de déterminer la valeur du quotient ouvert Q_o pour chaque cycle vibratoire. L'allure temporelle du signal *EGG* est une caractéristique utile pour distinguer les méca-

nismes I et II car la forme d'onde est plus asymétrique en MI qu'en MII [CL91, Hen01, HRC⁺03].

1.2 Aspects physiques : Pressions de seuil des oscillations

La modélisation physique du comportement mécanique des cordes vocales permet de prédire l'apparition et l'extinction du voisement, qui résulte d'un phénomène d'auto-oscillation. De nombreuses études physiques s'intéressent à l'influence des différents paramètres et phénomènes physiques sur la pressions de seuil des oscillations. Ces études sont soit des travaux théoriques basés sur des modèles théoriques décrivant le comportement des cordes vocales, soit des travaux expérimentaux réalisés en utilisant des larynx excisés ou bien des maquettes mécaniques auto-oscillantes.

La notion de pression de seuil des oscillations est définie puis ces travaux, s'intéressant à l'influence des paramètres géométriques, mécaniques et acoustiques sur la pression de seuil, sont résumés.

1.2.1 Définition

La notion de pression de seuil inclut deux grandeurs physiques différentes. La première est la pression P_{on} d'apparition (onset) des oscillations, définie comme la valeur minimum de pression sous-glottique nécessaire à l'apparition des oscillations entretenues. La deuxième grandeur caractérise la valeur de la pression sous-glottique pour le phénomène inverse. La pression d'extinction (offset) P_{off} est définie comme la valeur minimum de pression sous-glottique permettant d'entretenir les oscillations.

L'apparition et l'extinction de l'auto-oscillation lors de l'augmentation et la diminution de la pression sous-glottique sont deux phénomènes correspondant à une bifurcation du système laryngé. Une bifurcation implique qu'une petite variation d'un des paramètres physiques du système provoque une modification majeure de son comportement dynamique¹.

1.2.2 Hystérésis entre l'apparition et l'extinction des oscillations

Une des propriétés des oscillations induites par un écoulement [Luc99] est le phénomène d'hystérésis qui traduit que l'état du système lors de l'apparition des oscillations est différent de celui qui caractérise l'extinction des oscillations².

Ainsi, la pression sous-glottique P_{on} nécessaire à l'apparition des oscillations est toujours supérieure ou égale à la pression P_{off} à l'extinction des oscillations. Ce phénomène a été étudié analytiquement [Luc99, LKL⁺11], et observé dans des études réalisées sur des larynx de canins excisés [Bae81, RTZJ08] ou sur des maquettes de cordes vocales

1. Selon la théorie des systèmes dynamiques, l'apparition des oscillations correspond à une bifurcation de Hopf, qui traduit une transition d'un point d'équilibre fixe vers un cycle limite.

2. Ce comportement témoigne de la présence d'une bifurcation de Hopf sous-critique.

[TST95, CTT97, Rut07, RPH⁺07]. Le rapport P_{off}/P_{on} est généralement compris entre 0,5 et 1 [Bae81, TST95, CTT97]. Dans le cas de larynx canins, Regner et. al [RTZJ08] concluent que ces valeurs sont comprises entre 0,7 et 1.

1.2.3 Comportement de la pression P_{on}

Les paragraphes suivants décrivent l'influence des différents paramètres physiques sur l'évolution de la pression de seuil des oscillations P_{on} . Le comportement de la pression de seuil à l'extinction des oscillations P_{off} , qui est en première approximation équivalent à celui de P_{on} , ne sera pas mentionné.

Influence de la configuration glottique

De nombreuses études réalisées sur des maquettes de cordes vocales montrent que la valeur de P_{on} a tendance à augmenter avec l'écartement initial entre les cordes vocales, également appelé aperture glottique initiale [TST95, CT06]. Les études théoriques [FS11, Luc98] confirment ce comportement pour de grandes valeurs d'ouverture (ou d'aperture). Pour des valeurs plus faibles, les modèles qui tiennent compte de la viscosité de l'écoulement prédisent une forte augmentation de P_{on} lorsque l'aperture diminue [Luc98]. Une étude réalisée sur la maquette de cordes vocales présentée en section 4.2.1 et utilisée dans cette étude, montre un comportement plus complexe comprenant un maximum local de P_{on} pour une faible valeur d'aperture.

L'angle initial de convergence du canal glottique dans le sens de l'écoulement est également un paramètre qui influe sur l'apparition des oscillations. Chan et al. [CTT97] mesurent des valeurs de P_{on} minimales pour un angle de convergence nul tandis que Lucero [Luc98] prédit que la configuration optimale favorisant l'apparition des oscillations correspond à une glotte légèrement divergente. L'augmentation de l'angle glottique (angle formé par les axes des deux cordes vocales) produit une augmentation de P_{on} [PL12].

Influence des caractéristiques mécaniques des cordes vocales

Certaines études cherchent à caractériser l'effet de l'augmentation de la fréquence de la résonance mécanique des cordes vocales F_m sur la pression de seuil P_{on} . En augmentant le module d'Young du caoutchouc constituant leurs répliques de corde vocale, Zhang et al. [ZNB09] mesurent une augmentation de P_{on} . Sur des larynx excisés, le même effet se produit lorsque la tension longitudinale appliquée sur les cordes vocales [RTZJ08] est augmentée.

Dans cette étude, la maquette de cordes vocales utilisée est composée de deux tubes en latex remplis d'eau (cf. section 4.2.3). La modification de la tension des cordes vocales est imitée en modifiant la pression P_e de l'eau. Les mesures précédemment réalisées sur cette même maquette montrent que la relation décrivant l'évolution de la pression de seuil P_{on} en fonction de la pression P_e se présente sous la forme d'une courbe en « U » avec un minimum de

P_{on} pour $600 \leq P_e \leq 900 \text{ Pa}$ [PL12, Her14]. En utilisant une maquette dont le fonctionnement est similaire, Titze et al. [TST95] montrent que l'augmentation de la viscosité du fluide contenu dans les répliques de corde vocale, provoque également une augmentation de P_{on} . Ces résultats sont cohérents avec ceux de Cullen et al. [CGC00] qui prédisent, dans le cas du trombone, une diminution de P_{on} lorsque le facteur de qualité de la résonance mécanique augmente.

Influence des interactions acoustiques

Dans des conditions représentatives d'une production de parole en voix parlée, lorsque la fréquence fondamentale F_0 des oscillations est très inférieure à la fréquence F_1 de la première résonance du conduit vocal, l'augmentation de la longueur du conduit vocal a pour effet de diminuer la pression P_{on} ainsi que la fréquence F_0 [Rut07, LLH⁺12].

L'évolution de P_{on} est plus complexe lorsque l'intervalle entre F_0 et l'une des fréquences de résonance acoustique devient plus faible.

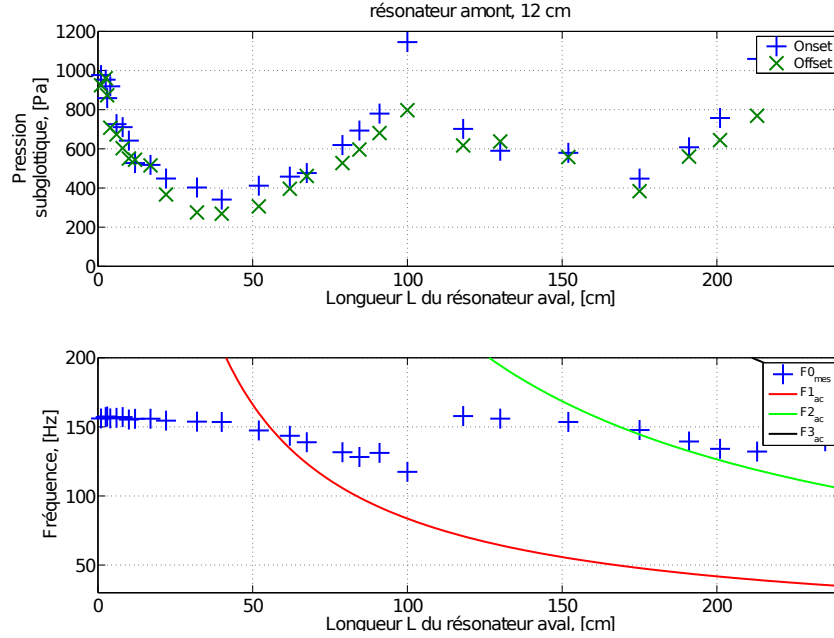


FIGURE 1.3 – Mesure des pressions P_{on} et P_{off} et de la fréquence fondamentale au seuil de l'oscillation en fonction de la longueur du résonateur acoustique situé en aval de la maquette de cordes vocales, d'après [Rut07]. Les courbes F_{iac} , pour $i = 1, 2, 3$, représentent la fréquence de la i ème résonance acoustique du résonateur aval.

L'augmentation de la fréquence F_m ou de la longueur du résonateur met en évidence l'existence de plusieurs régimes d'oscillation associés aux différentes résonances acoustiques, les transitions entre ces régimes se manifestent alors par des sauts de la fréquence fondamentale F_0 vers des valeurs plus importantes. Ce phénomène est observé sur maquette [CGC00, Rut07, ZNB09, LLH⁺12] et prédit en utilisant différents modèles mécaniques [CGC00, Rut07, LLH⁺12]. Dans ce cas de figure, l'évolution de P_{on} présente des minimums qui correspondent aux cas où la fréquence F_0 est légèrement inférieure à une fréquence

de résonance acoustique et des maximums locaux situés au niveau des transitions vers les régimes d'oscillations voisins. La figure 1.3 illustre ce phénomène.

1.3 Aspects phonétiques

Cette section a pour objectif de présenter les notions de phonétique utilisées dans la suite des travaux.

1.3.1 Voyelles

Une voyelle est phonétiquement définie comme un son du langage humain dont la production est caractérisée par la vibration des cordes vocales et à l'inverse des consonnes, sans aucune obstruction des voies aériennes supérieures. Les différentes voyelles d'une langue sont réalisées en modifiant la position des organes articulateurs. L'effet acoustique de ces ajustements de la géométrie du conduit vocal est, par effet de résonance, une modification de la répartition fréquentielle de l'énergie du signal émis par la source laryngée. La voyelle est dite « orale » lorsque les ondes acoustiques se propagent uniquement dans le conduit vocal et « nasale » lorsque le voile du palais est abaissé et que la cavité nasale participe aux résonances acoustiques.

Pour les voyelles orales, la répartition de l'énergie fréquentielle est caractérisée à l'aide des fréquences F_i de résonance du conduit vocal, également appelées formants, avec $i = 1$ pour la résonance la plus basse et $i = 2$ pour la résonance suivante. La figure 1.4 représente la répartition des voyelles orales du français dans l'espace formé par les fréquences F_1 et F_2 .

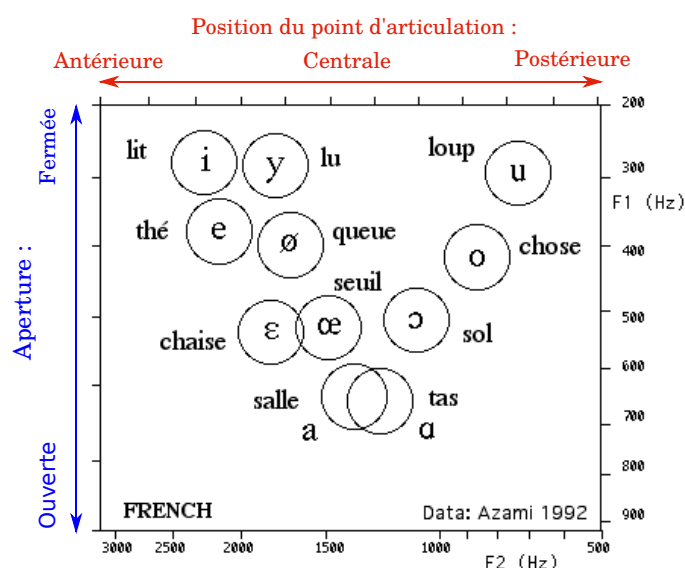


FIGURE 1.4 – Représentation des voyelles orales du français dans le plan (F_1, F_2) , d'après [Aza92]

La figure 1.4, appelée triangle vocalique, fait référence à la position de la langue à l'intérieur de la bouche. Les voyelles fermées (ou voyelles hautes) sont produites en réduisant l'écart entre le palais et la langue tandis que les voyelles ouvertes (ou voyelles basses) sont produites en abaissant la langue. Ainsi, l'aperture au point d'articulation influence la valeur de la fréquence F_1 . De même, la position du point d'articulation sur l'axe antero-postérieur influence la valeur de F_2 .

Les interactions acoustiques entre l'oscillation des cordes vocales et les résonateurs acoustiques sont significatives lorsque la fréquence fondamentale F_0 est proche d'une fréquence de résonance acoustique. Dans le cas de la voix parlée ($100 \leq F_0 \leq 250 \text{ Hz}$), cette condition correspond à la production de voyelles fermées pour lesquelles la fréquence F_1 est basse.

1.3.2 Fréquence fondamentale intrinsèque des voyelles

De nombreuses études phonétiques montrent qu'en moyenne, la fréquence fondamentale F_0 est plus importante pour les voyelles fermées que pour les voyelles ouvertes. Cette tendance, appelée fréquence fondamentale intrinsèque des voyelles (« Intrinsic pitch of the vowel ») est connue depuis la fin du 19^{ème} siècle [Mey97] et observée dans toutes les langues du monde [WL95]. Le caractère universelle de cet effet entraîne à penser que celui-ci n'est pas volontairement induit pour des raisons phonétiques, par exemple pour améliorer la distinction entre les différentes voyelles [WL95].

Lieberman [Lie70] suggère que l'interaction acoustique est à l'origine de ce phénomène. Le couplage acoustique plus important pour les voyelles hautes provoquerait une augmentation de la fréquence fondamentale de l'auto-oscillation. Cette hypothèse est contredite par Ohala et Eukel [OE76] et Ewan [EO79] qui montrent que cet effet persiste lorsque que les valeurs des fréquences de résonance acoustique sont constantes. De plus, les mesures sur maquette montrent que, avec des répliques de corde vocale dont les caractéristiques mécaniques sont fixées, la diminution de la valeur de F_1 provoque une baisse de la fréquence F_0 (cf. figure 1.3).

L'hypothèse mécanique suppose que lors de la production de voyelles fermées, la langue exerce indirectement une tension sur les cordes vocales, ce qui se traduit par une augmentation de la fréquence fondamentale [Leh70]. En mesurant les mouvements de l'os hyoïde, situé au-dessus du larynx, Honda [Hon83] confirme cette hypothèse. Lors de la production de voyelles fermées, la base de la langue est déplacée vers l'avant. Ce mouvement est en partie réalisé par la contraction de deux paires de muscles qui tirent l'os hyoïde vers l'avant, ce qui entraîne un basculement du cartilage thyroïde vers l'avant et donc une tension longitudinale au niveau des cordes vocales, sans modification de l'activité musculaire au niveau du larynx.

1.3.3 Consonnes

À l'inverse des voyelles, les consonnes se caractérisent par une obstruction du conduit vocal. Les différentes consonnes se distinguent par leur mode d'articulation, leur point d'articulation

et par leur voisement.

Mode d'articulation

Le mode d'articulation fait référence à l'obstruction : le conduit vocal est seulement partiellement obstrué pour les consonnes fricatives tandis que son occlusion est complète dans le cas des consonnes plosives (également appelées occlusives).

Point d'articulation

L'obstruction peut être réalisée à différents endroits du conduit vocal, avec les lèvres (bilabiale), avec la langue au niveau des dents (alvéolaire) ou au niveau de l'arrière palais (vélaire) pour les principaux.

Voisement

Le mode de voisement fait référence au fait que la production de certaines consonnes dites voisées (ou sonores) implique la vibration des cordes vocales, ce qui n'est pas le cas des consonnes dites non voisées (ou sourdes).

Consonnes plosives

Dans le cadre de cette thèse, l'intérêt est plus particulièrement porté sur les consonnes plosives. Pour les plosives orales, l'écoulement d'air provenant des poumons est complètement stoppé. Une surpression apparaît alors en amont du point de l'articulation. Au moment où l'occlusion est relâchée, un bruit aéroacoustique (explosion suivi d'une friction) est produit. Les plosives sont voisées dans le cas où les cordes vocales vibrent lors du relâchement de l'occlusion (et non voisées en l'absence de vibration laryngée à cet instant). La figure 1.5 illustre la position des articulateurs pour les différentes plosives.



FIGURE 1.5 – Représentation du conduit vocal lors de l'occlusion, pour la production des plosives orales de la langue française, schéma adapté de [Vai06]

1.3.4 Voice onset Time

Le *Voice Onset Time* (VOT) est un critère phonétique qui caractérise la production de plosives. Le VOT est défini par l'intervalle de temps entre le relâchement de l'occlusion au point d'articulation et l'apparition du voisement. L'instant de relâchement de l'occlusion peut être déterminé en repérant l'effet de l'explosion aérodynamique sur le signal acoustique ou l'instant de l'ouverture de l'occlusion à partir d'enregistrements vidéos. La définition de

l'instant correspondant à l'apparition des oscillations est plus floue. De nombreux auteurs utilisent la définition de Klatt [Kla75], qui mesure sur le spectrogramme du signal acoustique, l'apparition des stries verticales au niveau du deuxième formant et des formants supérieurs. Lisker et al. [LLDE75] déterminent le VOT à partir du moment où le premier formant atteint son amplitude maximale.

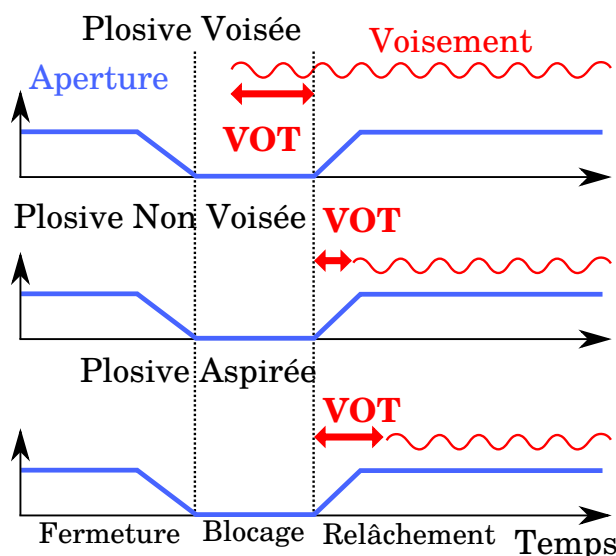


FIGURE 1.6 – Illustration de la notion de *Voice Onset Time* (VOT) dans le cas de la production de plosives voisées, non voisées et aspirées.

Le VOT est la caractéristique permettant de différencier les consonnes plosives produites avec le même mouvement articulaire. Le VOT est négatif dans le cas des plosives voisées (/b/, /d/, /g/), positif et faible pour les plosives non voisées (/p/, /t/, /k/), positif et relativement long dans le cas des plosives aspirées. Ces différents cas de production sont illustrés en figure 1.6.

1.4 Éléments sur la production de parole chez les sourds

Cette section résume les principaux troubles de la production de la parole observés chez les locuteurs souffrant de perte auditive sévère (sourds) ou moyenne (mal-entendants). Elle est inspirée des écrits de Osberger et McGarr [OM82] et de Giroux [Gir12] qui récapitulent les résultats de nombreuses études sur le sujet. Cette thèse s'intéresse aux effets de la surdité sur la production de la parole. Les troubles inhérents aux productions de locuteurs sourds pré et post-linguistique sont reportés ici sans distinction particulière. Selon Nickerson [Nic75], les trois problèmes les plus souvent observés dans la parole chez les sourds sont le manque de contrôle de la fréquence fondamentale, la nasalité trop présente et la voix trop soufflée.

1.4.1 Fréquence fondamentale

Valeur moyenne

Pour les adultes, la fréquence fondamentale moyenne produite par des locuteurs mal-entendants est généralement plus élevée que celle mesurée chez des locuteurs normo-entendants [AKH64, Boo66, Mar68]. Cette tendance n'est cependant pas observée pour tous les groupes d'individus, Osberger [Osb81] montre que l'effet inverse se produit parfois chez des sujets masculins, âgés de 13 à 15 ans. De plus, cette tendance n'est pas statistiquement significative pour des enfants âgés de 6 à 12 ans [Gre56, Boo66, MEV79].

Une étude plus récente réalisée chez les enfants, dont l'âge moyen est de 9 ans, montre que ce phénomène est lié aux capacités auditives des sujets [RLG03]. En effet, les auteurs ne constatent aucune différence significative entre les groupes de locuteurs mal-entendants et normo-entendants mais ils observent que la fréquence fondamentale moyenne est plus élevée chez les sourds profonds.

Variation

Les variations de la fréquence fondamentale au cours d'une même phrase sont généralement moins importantes chez les mal-entendants [Voe35, Gre56, Hoo66] que chez les normo-entendants. De même que pour la fréquence fondamentale moyenne, cette tendance n'est pas confirmée pour tous les individus : certains sujets mal-entendants produisent des variations de fréquence fondamentale jugées excessives, de plus de 100 Hz au cours de la même phrase [MEV79, Smi75, SNR78]. La fréquence fondamentale varie également avec le contexte vocalique. Bush [Bus81] a montré que pour les sujets sourds capables de produire une large gamme de voyelles (c'est-à-dire reproduire un espace vocalique dans le plan (F_1, F_2) aussi important que celui des locuteurs normo-entendants), les voyelles fermées sont produites avec une fréquence fondamentale plus élevée que celle associée aux voyelles ouvertes. Cet effet qui correspond à la « fréquence fondamentale intrinsèque des voyelles » (cf. section 1.3.2, est plus marqué chez les sujets sourds : les variations de F_0 sont plus importantes. L'auteur attribue ce résultat au fait que la fréquence fondamentale, et donc la tension des cordes vocales, est en moyenne plus importante chez les sourds. La caractéristique non linéaire de la relation contrainte-déformation des cordes vocales implique que la tension supplémentaire induite par la production d'une voyelle fermée serait à l'origine d'une augmentation de F_0 plus importante.

1.4.2 Nasalisation

Le contrôle de l'abaissement et de l'élévation du voile du palais, permettant d'ouvrir et de fermer la cavité nasale, est reconnu pour être une source de difficulté importante pour les personnes sourdes et mal-entendantes. La nasalisation excessive a été reportée dans de nombreuses études pour la production de voyelles [Mar68] et de consonnes [HN42, Smi75, SNBR76]. À partir de l'amplitude des vibrations mesurées à l'aide d'un accéléromètre fixé sur le nez des sujets, Stevens et al. [SNBR76] reportent que 76 % des enfants sourds profonds nasalisent

excessivement pour au moins la moitié des voyelles prononcées. Ce phénomène apparaît également pour les consonnes. Les plosives orales par exemple sont souvent substituées par des plosives nasales (/m/, /n/ et /ɲ/) [HN42].

1.4.3 Voyelles

De nombreuses études montrent que, chez les sujets sourds, la production des voyelles est caractérisée par une réduction de l'espace vocalique, les fréquences des formants F_1 et F_2 tendent vers celles de la voyelle neutre /ə/ [OLS79, Bus81, MH80, Mai11, Mon78, AKH64]. Cette réduction de l'espace vocalique peut s'expliquer en partie par la restriction des mouvements de la langue, en particulier dans la direction antéro-postérieure, ce qui restreint l'étendue des valeurs de la fréquence F_2 . Cette observation peut s'expliquer par le fait que la position avant-arrière de la langue n'est pas bien visible par les sujets sourds, alors qu'une partie de leur apprentissage est basé sur l'imitation des gestes articulatoires observés [OM82, MG83].

1.4.4 Consonnes

Erreurs liées au voisement

Selon Hudgins et Numbers [HN42], l'erreur la plus fréquemment produite lors de la prononciation de consonnes est la confusion entre une consonne voisée et son homologue non voisée (produite avec le même mode d'articulation et au même endroit mais sans vibration des cordes vocales). Les études phonétiques montrent qu'il n'y a pas de direction privilégiée pour cette substitution. Certaines études reportent que la substitution est le plus souvent produite en faveur de la consonne voisée [Smi75] et d'autre reportent que la substitution vers l'homologue non voisée est plus courante [Nob67].

Erreurs liées au point d'articulation

Les erreurs de substitution commises sur le lieu d'articulation sont fortement corrélées avec la visibilité des articulateurs [Nob67, Smi75, Gol78]. Moins les articulateurs sont visibles et plus les erreurs commises pour prononcer les consonnes plosives correspondantes sont fréquentes. Ainsi, les consonnes bilabiales sont les mieux reproduites. D'autre part, le lieu d'articulation des consonnes plosives et des fricatives est moins différencié chez les sourds que chez les locuteurs normo-entendants [LW91].

Autres troubles

Des mesures aérodynamiques [HCS94] (pression intra-orale et débit nasal) suggèrent que, pour les mal-entendants, les valeurs de pression sous-glottique et intra-orale sont considérablement plus importantes que celles mesurées chez les normo-entendants. L'absence de retour auditif affecte également les caractéristiques temporelles de la production, Waldstein [Wal90]

note que, chez les sourds, le débit de parole est plus lent que chez les normo-entendants, en raison notamment d'une durée plus importante des voyelles. Le timbre de la voix peut également être altéré. Lorsque le niveau d'adduction entre les cordes vocales est trop faible, les effets de turbulence au niveau de la constriction glottique génèrent un bruit aéroacoustique large bande. Cette voix soufflée est notamment produite durant la production d'un son voisé, dans le cas où le débit d'air est trop important. L'accolement partiel entre les cordes vocales durant la vibration engendre un signal de débit glottique ne contenant que peu d'énergie en haute fréquence [SNR78].

1.5 Bilan

La production de son voisés est régie par l'auto-oscillation des cordes vocales. Le niveau de pression sous-glottique nécessaire à l'apparition de ce phénomène varie en fonction des ajustements laryngés et articulatoires réalisés par le locuteur. L'utilisation de modèles mécaniques simplifiés des cordes vocales permet de prédire l'évolution globale du niveau de pression sous-glottique au seuil des oscillations. D'autre part, ce phénomène de voisement est impliqué dans la majorité des sons qui constituent une langue.

Le récapitulatif des études sur la production de parole chez les sourds fait apparaître que la nature des troubles observés est variée et qu'il y a une grande variabilité inter-individuelle entre les différents sujets. Seules les stratégies de production impliquant des articulateurs visibles sont correctement réalisées. Ce résultat peut s'expliquer par l'importance du *feedback* visuel dans la construction des modèles internes chez les locuteurs sourds.

Ces travaux de thèse s'inscrivent dans le cadre des troubles de la production liés à un défaut de contrôle des aspects fréquentiels et temporels du voisement, tels que des variations excessives de la fréquence fondamentale et des substitutions de plosives voisées par des plosives non voisées (et inversement). Pour ce type de troubles de la production, le comportement dynamique de la source laryngée peut être prédit en utilisant des modèles mécaniques simplifiés de cordes vocales.

Modèles théoriques et implémentation numérique

Sommaire

2.1	Introduction	23
2.2	Modèles aérodynamiques	24
2.2.1	Approximations aérodynamiques	24
2.2.2	Écoulement quasi-stationnaire non-visqueux	27
2.2.3	Prise en compte de la viscosité	27
2.2.4	Séparation de l'écoulement	28
2.2.5	Application au modèle à deux masses	29
2.3	Modèles mécaniques	31
2.3.1	Modèle à deux masses de cordes vocales	31
2.3.2	Expression de la force de pression	32
2.3.3	Expression de la force visqueuse	32
2.3.4	Expression de la force élastique	33
2.3.5	Modélisation de la collision	34
2.4	Modèles acoustiques	35
2.4.1	Modèle de propagation	36
2.4.2	Pertes viscothermiques	40
2.4.3	Modèle de rayonnement	45
2.4.4	Impédance d'entrée des résonateurs acoustiques	47
2.4.5	Couplage acoustique-aérodynamique	48
2.5	Implémentation numérique	49
2.6	Bilan	51

2.1 Introduction

Ce chapitre présente les modèles théoriques utilisés pour les simulations numériques. Les hypothèses et les équations relatifs aux différents domaines physiques sont présentées. L'organisation du code de simulation est ensuite expliquée.

La première approche utilisée pour modéliser la structure mécanique des cordes vocales est basée sur l'hypothèse d'éléments localisés. Cette approximation, valable en basse fréquence

suppose que la dimension caractéristique de la structure mécanique est très petite devant la longueur d'onde des phénomènes considérés. En considérant que la longueur caractéristique des cordes vocales est de l'ordre du centimètre, l'approximation des éléments localisés est valable pour des fréquences inférieures à 10 kHz . Les éléments localisés sont souvent associés à l'hypothèse d'une déformation limitée à une seule direction. Un exemple représentatif de ce type de modélisation est le modèle à une masse de Flanagan et Landgraf [FL68].

Plus récemment, l'augmentation des capacités de calculs des machines a fait émerger l'utilisation de méthodes de résolution numérique tel que les éléments finis [ABT00, TJZ06], qui permettent l'utilisation de modèles mécaniques continus pour prédire le mouvement des cordes vocales. Ce type de modèle permet notamment de prendre en compte une géométrie plus réaliste en trois dimensions et d'améliorer la résolution spatiale de la simulation.

Entre ces deux extrêmes, plusieurs configurations de structure mécanique incluant plusieurs masses ont été proposées pour représenter le comportement mécanique des cordes vocales. Parmi ces modèles de type éléments distribués, le plus populaire est le modèle à deux masses de Ishikaza et Flanagan [IF72], qui se présente comme un compromis intéressant entre la simplicité et le réalisme de la modélisation. D'après ces auteurs, deux masses sont suffisantes pour simuler le déphasage observé entre les mouvements vibratoires des extrémités amont et aval de la glotte et reproduire ainsi l'évolution de l'aire glottique. Pelorson et al. [PHvH⁺94] proposent d'exploiter les positions des deux masses afin de raffiner la géométrie de la constriction glottique et la modélisation de l'écoulement glottique. Dans la même optique, Lous et al. [LHVH98] proposent un modèle à deux masses dans lequel l'évolution de la hauteur de la constriction glottique dans le sens de l'écoulement est linéaire par morceaux.

Le modèle utilisé dans le cadre de ce travail est basé sur le modèle à deux masses de Lous. L'avantage de ce type de modèle simplifié est d'une part le faible nombre de paramètres de contrôle et le coût de calcul relativement bas. Ces modèles sont notamment adaptés à une implémentation en temps réel. L'inconvénient par rapport à des modèles plus réalistes est la difficulté à établir le lien entre les paramètres du modèle mécanique et les caractéristiques physiologiques des tissus humains ou des matériaux qui constituent les répliques de corde vocale.

2.2 Modèles aérodynamiques

Cette section présente le modèle théorique utilisé dans les simulations numériques pour décrire l'écoulement de l'air à travers la glotte durant la phonation.

2.2.1 Approximations aérodynamiques

Le mouvement de l'air est régi par les équations de Navier-Stokes qui traduisent la conservation de la masse, de l'énergie et de la quantité de mouvement d'une particule de fluide. Les hypothèses émises sur la nature de l'écoulement d'air à travers le larynx, afin d'obtenir une

expression analytique, sont ici brièvement présentées.

Dimensions et grandeurs caractéristiques

Pour l'écoulement d'air, les cordes vocales forment une constriction dont la géométrie peut être représentée par le schéma de la figure 2.1. Le vecteur \vec{e}_x définit la direction principale de l'écoulement. La dimension caractéristique de la constriction, associée à cette direction est la longueur ℓ . Les directions transverses selon les vecteurs \vec{e}_y et \vec{e}_z sont associées respectivement à l'aperture h et à la largeur L de la constriction.

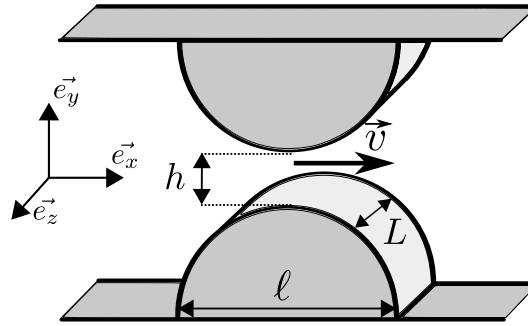


FIGURE 2.1 – Dimensions caractéristiques de la constriction glottique.

Corps humain	Repère cartésien	Dimensions	Ordre de grandeur
Axe vertical	\vec{e}_x	ℓ	$2 - 4 \text{ mm}$
Axe latéral	\vec{e}_y	h	$0 - 1 \text{ mm}$
Axe antero-postérieur	\vec{e}_z	L	$1,5 - 2,5 \text{ mm}$

TABLE 2.1 – Dimensions typiques de la constriction glottique et orientation dans le corps humain.

Approximations générales

Pour l'écoulement glottique, les approximations suivantes peuvent être raisonnablement effectuées [Vil02] :

- l'air peut être considéré comme un fluide Newtonien,
- l'effet de la force gravitationnelle sur l'écoulement est négligeable,
- l'écoulement peut être considéré comme incompressible,
- l'écoulement d'air est adiabatique,
- l'effet Coanda n'a pas le temps de s'installer durant une période de vibration glottique (pour $F_0 = 100 \text{ Hz}$) [PHvH⁺94, Hof98].

Ces approximations seront adoptées tout au long de cette section.

Équations de Navier-Stokes

Sous les hypothèses mentionnées ci-dessus, les équations de Navier-Stokes se résument au système suivant :

$$\nabla \cdot \mathbf{v} = 0, \quad (2.1)$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\rho} \nabla P + \nu \Delta \mathbf{v}, \quad (2.2)$$

où P représente la pression aérodynamique, \mathbf{v} désigne la vitesse de l'écoulement, qui s'exprime dans un repère cartésien telle que $\mathbf{v} = v_x \vec{e}_x + v_y \vec{e}_y + v_z \vec{e}_z$ et ν le coefficient de viscosité cinématique de l'air.

Quasi-stationnarité

Le nombre de Strouhal Sr est défini comme le rapport entre les termes d'accélération instationnaire et d'accélération convective de la particule de fluide, tel que $Sr = |\frac{\partial \mathbf{v}}{(\mathbf{v} \cdot \nabla) \mathbf{v}}| \sim F_0 \ell / v$, où $v = |\mathbf{v}|$. Le caractère quasi-stationnaire de l'écoulement est vérifié pour une production typique en mécanisme I, avec $F_0 \approx 100 \text{ Hz}$ et $v \approx 40 \text{ m.s}^{-1}$, le nombre de Strouhal correspondant est alors $Sr \approx 10^{-2}$. Cette hypothèse est moins évidente dans le cas du mécanisme laryngé II et lorsque la pression sous-glottique est faible : en considérant une fréquence fondamentale $F_0 \approx 400 \text{ Hz}$ et une vitesse $v \approx 15 \text{ m.s}^{-1}$, la valeur de nombre de Strouhal devient $Sr \approx 10^{-1}$. Dans les deux cas, l'écoulement devient instationnaire lorsque l'aperture glottique h tend vers zéro, dans le cas d'une fermeture complète de la glotte.

Écoulement laminaire

Le nombre de Reynolds Re représente le rapport des forces inertielles sur les forces visqueuses. Il est défini tel que $Re = vd/\nu$, où d est la dimension caractéristique de l'obstacle ou de la conduite. Dans le cas de la constriction glottique, il est judicieux d'utiliser le nombre de Reynolds Re_h basé sur la plus petite dimension de la constriction tel que $Re_h = vh/\nu$.

La dimension longitudinale ℓ est également importante : le nombre de Reynolds $Re_\ell = v\ell/\nu$ basé sur la dimension ℓ donne aussi une information sur l'apparition des turbulence dans la glotte. Pour un conduit lisse, un écoulement ne se séparant pas de la paroi peut être considéré comme laminaire lorsque $Re_\ell < 10^5$. Cette hypothèse est vérifiée dans le cas de l'écoulement glottique ($Re_\ell \sim 10^4$ en prenant $\nu = 1.5 \cdot 10^{-5} \text{ m}^2.\text{s}^{-1}$).

Ces deux dimensions doivent cependant être prises en compte pour obtenir une bonne indication du comportement de l'écoulement dans la constriction glottique, il est alors possible d'utiliser la quantité $Re_h h / \ell = v h^2 / (\nu \ell)$ [DPV+03]. Le comportement du fluide est dominé par la viscosité si $Re_h h / \ell < 1$. Dans la glotte, l'ordre de grandeur de la quantité $Re_h h / \ell$ est de 10^2 . En pratique, le comportement du fluide dépendent de la géométrie de la constriction, l'écoulement glottique peut être considéré comme laminaire uniquement dans la partie convergente de la glotte.

2.2.2 Écoulement quasi-stationnaire non-visqueux

En considérant un écoulement quasi-stationnaire pour lequel les forces visqueuses sont négligeables (approximation de fluide parfait pour $Re \rightarrow \infty$), les équations régissant le mouvement d'une particule de fluide s'écrivent :

$$\nabla \cdot \mathbf{v} = 0, \quad (2.3)$$

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\rho} \nabla P. \quad (2.4)$$

Équation de Bernoulli stationnaire

L'intégration de l'équation 2.4 le long d'une ligne de courant pour un écoulement irrotationnel permet d'obtenir l'équation de Bernoulli stationnaire :

$$P(x) + \rho \frac{v^2(x)}{2} = cte. \quad (2.5)$$

L'équation de Bernoulli traduit la conservation de l'énergie mécanique qui est définie comme la somme de l'énergie potentielle et cinétique. Celle-ci ne s'applique donc pas aux écoulements dissipatifs.

2.2.3 Prise en compte de la viscosité

Écoulement de Stokes

Les effets induits par la viscosité sont traités en considérant un écoulement quasi-stationnaire dont la dynamique est dominée par les effet visqueux et pour lequel les effets inertiels sont négligeables ($Re \ll 1$). Sous ces hypothèses, l'équation 2.2, traduisant la conservation de la quantité de mouvement devient :

$$\Delta \mathbf{v} = \frac{1}{\mu} \nabla P \quad (2.6)$$

où μ est la viscosité dynamique de l'air telle que $\mu = \nu \rho$.

Écoulement de Poiseuille

Pour un écoulement unidimensionnel (cf. Annexe A) tel que $\mathbf{v} = v_x(y) \vec{e}_x$, dans une constriction de hauteur h uniforme, les conditions de non-glissement au niveau des parois ($v_x|_{y=0} = 0$ et $v_x|_{y=h} = 0$) impliquent que le profil de vitesse s'exprime :

$$v_x(y) = \frac{1}{2\mu} \left(\frac{\partial P}{\partial x} \right) (y - h)y. \quad (2.7)$$

Expression du gradient de pression

Pour un écoulement incompressible et quasi-stationnaire, l'équation de continuité qui permet de relier le débit volumique U à la vitesse v_x s'écrit :

$$U = L \int_{y=0}^h v_x(y) dy. \quad (2.8)$$

En remplaçant $v_x(y)$ par l'expression du profil de vitesse d'un écoulement de Poiseuille (équation 2.7), l'équation de continuité permet d'exprimer le gradient de pression en fonction du débit volumique et de la hauteur de la constriction, tel que :

$$\frac{\partial P}{\partial x} = \frac{12\mu U}{L} \frac{1}{h^3}. \quad (2.9)$$

2.2.4 Séparation de l'écoulement

Description

Le modèle d'écoulement utilisé prends en compte l'apparition d'une éventuelle séparation de l'écoulement, dans le cas d'une géométrie divergente, en aval du point pour lequel la hauteur de la constriction est minimale. La séparation de l'écoulement est due aux interactions entre les forces d'inerties et les forces visqueuses. Les phénomènes physiques responsables de la séparation sont présentés en annexe A. Pour la glotte, le phénomène de séparation de l'écoulement se traduit par la formation d'un jet en aval de la constriction. Une grande partie de l'énergie cinétique du jet est dissipée par le phénomène de turbulence qui apparaît à l'interface entre la zone pariétale où la vitesse de l'écoulement est nulle et la zone principale associée au jet. Il est alors raisonnable de considérer que la pression moyenne en aval est égale à la pression atmosphérique.

Modélisation

Le modèle de séparation de l'écoulement utilisé est le critère semi-empirique de Liljencrants, utilisé dans de nombreuses études [LHVH98, Luc98, Luc99, Vil02, SD04, LK05, Rut07, Bai09]. Ce critère repose sur l'hypothèse selon laquelle la position du point de séparation dépend uniquement de la géométrie de la constriction et n'est pas influencée par les caractéristiques de l'écoulement. Il se base sur des études théoriques et expérimentales [PHvH⁺94, PVC⁺96]

qui montrent que, dans le cas d'une glotte divergente, la relation entre le minimum de l'aire glottique A_{min} et l'aire au point de séparation A_s tend vers une valeur constante lorsque $Re \rightarrow 10^3$, telle que :

$$A_s = \alpha A_{min}. \quad (2.10)$$

Ce critère n'a de sens que lorsque l'écoulement peut être considéré comme stationnaire. Les valeurs couramment utilisées ou suggérées pour le coefficient α sont 1,1 [PVC⁺96], 1,2 [LHVH98] et 1,3 [Luc99].

Des simulations numériques basées sur la théorie des couches limites (cf. Annexe A) permettent d'estimer la validité de ce critère. Ainsi, il apparaît que le rapport α diminue considérablement avec l'augmentation de l'angle de divergence de la constriction [Hof98] et qu'il augmente lorsque la hauteur h de la constriction diminue et ℓ augmente [Cis08, HCP09].

Influence sur la dynamique de l'auto-oscillation

Dans le modèle à deux masses de Ishikaza et Flanagan [IF72], le point de séparation est implicitement fixé à l'extrémité du canal glottique formé par les deux masses. Cependant, des études plus récentes ont montré que la position du point de séparation de l'écoulement joue un rôle essentiel dans le transfert d'énergie entre l'écoulement et le mouvement des cordes vocales, et donc dans le mécanisme d'auto-oscillation.

Dans le cas d'une glotte divergente, la prise en compte d'un point de séparation mobile le long de la constriction glottique modifie le comportement dynamique du système. La fréquence fondamentale diminue ainsi de 15 % et l'allure du débit glottique simulé est significativement modifiée [PHvH⁺94]. Selon Lucero [Luc98], dans le cas d'une constriction glottique divergente et relativement large, les effets de la séparation de l'écoulement sont bien plus importants que les effets liés à la viscosité.

La position du point de séparation a également une influence directe sur le rapport P_{off}/P_{on} [Luc99]. Les simulations réalisées avec le modèle de Ishikaza et Flanagan [IF72], en considérant un point de séparation localisé à la jonction entre les deux masses ($\alpha = 1$) dans le cas où la glotte est divergente [HK95, SH95, Sto95], ne peuvent pas reproduire le phénomène d'hystérésis ($P_{off}/P_{on} < 1$) qui caractérise le système laryngé (cf. section 1.2).

2.2.5 Application au modèle à deux masses

Géométrie de la constriction glottique

Dans le modèle à deux masses utilisé, la glotte est représentée par une constriction dont la hauteur $h(x)$ varie linéairement par morceaux suivant x . La géométrie de cette constriction est décrite par quatre points fixes d'abscisse x_i pour $i = 0, 1, 2, 3$. Entre chacun de ces points, la hauteur de la constriction est donnée par :

$$h(x, t) = \frac{h_i(t) - h_{i-1}(t)}{(x_i - x_{i-1})}(x - x_{i-1}) + h_{i-1}(t), \text{ pour } x_{i-1} \leq x \leq x_i. \quad (2.11)$$

Les hauteurs h_0 et h_3 sont constantes. La figure 2.2 illustre cette géométrie.

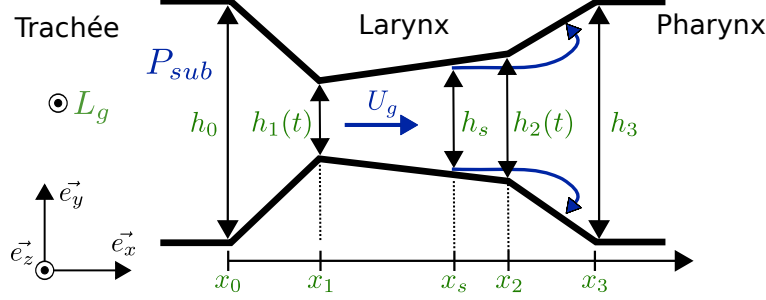


FIGURE 2.2 – Géométrie de la constriction glottique associée au modèle à deux masses de Lous [LHVH98] dans le cas d'une constriction divergente. Les variables x_s et h_s représentent respectivement l'abscisse et l'aperture au point de séparation de l'écoulement.

Le point de séparation de l'écoulement est autorisé à se déplacer entre les deux masses et l'aperture h_s au niveau du point de séparation, pour $x = x_s$, est donnée par la condition suivante $h_s = \min(\alpha h_1, h_2)$. Ces deux conditions imposent que :

$$\begin{aligned} h_s &= \alpha h_1 & \text{et} & \quad x_1 < x_s \leq x_2, & \text{si } \alpha h_1 < h_2, \\ h_s &= h_2 & \text{et} & \quad x_s = x_2, & \text{si } \alpha h_1 \geq h_2. \end{aligned} \quad (2.12)$$

Équation de Bernoulli corrigée pour la viscosité

Ce modèle d'écoulement unidimensionnel et quasi-stationnaire correspond à l'équation de Bernoulli (équation 2.5) à laquelle est ajouté un terme correctif provenant de l'expression du gradient de pression pour un écoulement de Poiseuille (équation 2.9), afin de tenir compte des pertes visqueuses.

Pour une constriction de largeur L_g constante dont la hauteur $h(x)$ varie dans la direction principale de l'écoulement, la chute de pression ΔP_{visc} due à la viscosité peut être déterminée en intégrant l'expression donnée par l'équation 2.9. La chute de pression s'exprime alors :

$$\Delta P_{visc}(x, t) = \frac{12\mu U_g(t)}{L_g} \int_0^x \frac{1}{h^3(x, t)} dx. \quad (2.13)$$

L'équation de Bernoulli corrigée, appliquée entre le point d'abscisse x_0 et le point de séparation de l'écoulement x_s s'écrit :

$$P(x_0, t) + \frac{\rho}{2} \frac{U_g^2(t)}{L_g^2 h_0(t)} = P(x, t) + \frac{\rho}{2} \frac{U_g^2(t)}{L_g^2 h^2(x, t)} + \Delta P_{visc}(x, t). \quad (2.14)$$

Cette relation permet de déterminer le débit glottique U_g en fonction de la valeur de la pression sous-glottique P_{sub} et de la géométrie $h(x)$ de la constriction glottique. En considérant que $P(x = x_0) = P_{sub}$ et que $P(x = x_s) = 0$, l'équation 2.14 devient :

$$P_{sub}(t) = \frac{\rho}{2} \frac{U_g^2(t)}{L_g^2} \left(\frac{1}{h_s^2(t)} - \frac{1}{h_0^2(t)} \right) + \Delta P_{visc}(x_s, t). \quad (2.15)$$

2.3 Modèles mécaniques

Les cordes vocales humaines sont constituées de plusieurs couches ayant chacune différentes caractéristiques mécaniques. Cette structure, dont les propriétés mécaniques sont hétérogènes et anisotropes, présente un comportement mécanique viscoélastique complexe et non linéaire. Le modèle mécanique général utilisé repose sur les hypothèses selon lesquelles la loi contrainte-déformation ne dépend ni des sollicitations passées ni de la fréquence de l'excitation ou de la température.

2.3.1 Modèle à deux masses de cordes vocales

Le modèle mécanique de cordes vocales utilisé est basé sur le modèle proposé par Lous et al. [LHVH98]. Il s'agit d'un modèle symétrique dans lequel chaque corde vocale est représentée par deux masses. Le déplacement des masses est limité à l'axe gauche-droite du corps humain (perpendiculaire à la direction de l'écoulement). La figure 2.3 illustre ce modèle.

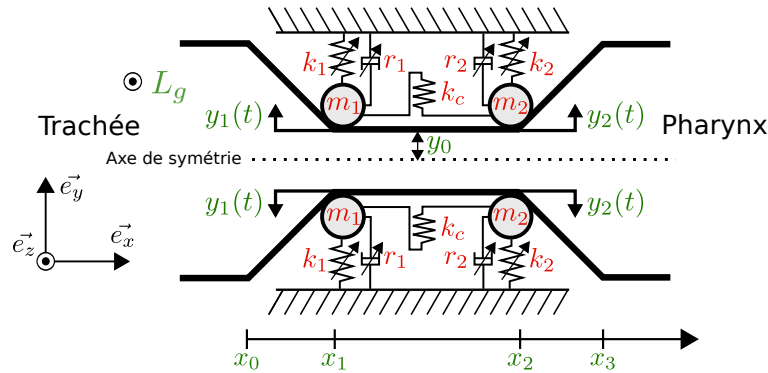


FIGURE 2.3 – Schéma du modèle mécanique symétrique de cordes vocales. Le sens de l'écoulement de l'air provenant des poumons est donné par le vecteur \vec{e}_x .

La position d'équilibre des masses est repérée par la distance y_0 qui les sépare de l'axe de symétrie de la constriction glottique. La variable y_0 représente ainsi la demi-aperture de

la glotte au repos. L'équation du mouvement des masses m_i est donnée en fonction de leur déplacement y_i par rapport à leur position d'équilibre :

$$m_i \ddot{y}_i + F_{ri}(y_i, \dot{y}_i) + F_{ki}(y_i) + k_c(y_i - y_j) = F_i, \quad (2.16)$$

avec $i = 1, 2$ et $i \neq j$, où la notation \dot{y}_i représente la dérivée temporelle de la quantité y_i . Les notations k_c , F_i , F_{ri} et F_{ki} désignent respectivement la raideur de couplage du ressort idéal reliant les masses m_1 et m_2 , les forces de pression, visqueuse et élastique s'exerçant sur chacune des masses. Le lien entre le déplacement y_i et la hauteur h_i de la constriction glottique (appelée également ouverture glottique) au point d'abscisse x_i est donné par la relation $h_i = 2(y_i + y_0)$.

2.3.2 Expression de la force de pression

La force de pression F_i qui s'applique sur la masse m_i est l'intégrale de la distribution de pression $P(x)$ sur les parois adjacentes de la constriction glottique, pondérée par un coefficient tenant compte de la distance par rapport à l'abscisse de la masse considérée, telle que :

$$F_i(t) = L_g \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} P(x, t) dx + L_g \int_{x_i}^{x_{i+1}} \frac{x_{i+1} - x}{x_{i+1} - x_i} P(x, t) dx. \quad (2.17)$$

2.3.3 Expression de la force visqueuse

La dissipation de l'énergie mécanique est modélisée par la force F_{ri} . L'expression utilisée fait apparaître une non-linéarité qui ne dépend que de la valeur instantanée du déplacement y_i , telle que :

$$F_{ri}(y_i, \dot{y}_i) = r_i(1 + \eta_r \dot{y}_i^2) \dot{y}_i, \quad (2.18)$$

où r_i représente le coefficient d'amortissement visqueux et η_r l'amplitude de la contribution non linéaire de F_{ri} . En pratique, la valeur du coefficient r_i est souvent fixée en utilisant le taux d'amortissement relatif ξ défini par $\xi_i = r_i / (2\sqrt{k_i m_i})$. La valeur typique utilisée pour le modèle à deux masses est $\xi = 0,1$ [IF72, LHVH98]. Dans le cas où $\eta_r = 0$, l'expression de la force visqueuse correspond à celle utilisée par Ishikawa et Flanagan [IF72]. L'ajout d'une non-linéarité dépendante du déplacement a été proposé par Laje et al. [LGM01] pour le « flapping model » de Titze [Tit88b] puis adapté au modèle à deux masses [LK05]. La formulation utilisée ici, présentant une dépendance quadratique, est celle proposée par Lucero [Luc05] et reprise par Assaneo et Trevisan [AT13].

L'introduction de ce terme non linéaire se justifie par les données expérimentales issues de mesures sur des larynx excisés de canins [AT85]. La constante de temps, qui caractérise le retour à la position d'équilibre, augmente avec le niveau de déformation des tissus. Cette observation conduit à utiliser un modèle dans lequel l'amortissement augmente avec

le déplacement par rapport à la position d'équilibre. L'introduction de ce terme peut se justifier également par rapport au comportement des simulations. En effet, la simulation d'un mouvement d'abduction (écartement des cordes vocales) réalisé durant la phonation, en utilisant le modèle à deux masses de Ishikaza et Flanagan [IF72], prédit une augmentation de l'amplitude des oscillations du signal de débit glottique, or ce phénomène n'est pas observé expérimentalement [MKL95, LK05]. Cette non-linéarité dans l'expression de l'amortissement permet de limiter l'amplitude des oscillations du signal de débit glottique [LK05] sans avoir recours à un paramètre de contrôle supplémentaire [MKL95].

2.3.4 Expression de la force élastique

L'expression de la force élastique F_{ki} exercée par le ressort de raideur k_i sur la masse m_i s'écrit :

$$F_{ki}(y_i) = k_i y_i (1 + \eta_k y_i^2), \quad (2.19)$$

où k_i est la raideur des cordes vocales et le paramètre η_k l'amplitude de la contribution non linéaire de la force élastique par rapport au déplacement. Bien qu'absent du modèle de Lous [LHVH98], le terme non linéaire de cette expression, initialement introduit par Ishikaza et Flanagan [IF72], est repris dans les modèles mécaniques de cordes vocales plus récents [LK05, AT13].

L'introduction de ce terme non linéaire se justifie par la relation qui existe entre la pression sous-glottique et la fréquence fondamentale de la phonation. En effet, de nombreuses mesures *in vivo* [Bae79, LKM69] ou sur des larynx excisés de canins [Tit89, AS07] ou d'humain [Gar07] indiquent que les variations de pression sous-glottique sont positivement corrélées avec la fréquence fondamentale. Durant la phonation, une augmentation soudaine de pression sous-glottique produit immédiatement une augmentation de fréquence fondamentale, alors que les muscles thyro-aryténoïdien et inter-aryténoïdien sont encore inactifs [Bae79]. En approximant cette relation par une fonction affine, Baer relève des variations de l'ordre de $3 \cdot 10^{-2} \text{ Hz} \cdot \text{Pa}^{-1}$. Selon Alipour et Scherer [AS07], la relation entre ces deux paramètres est non linéaire et très influencée par le niveau d'abduction et l'élongation des cordes vocales.

De nombreuses études montrent que cette dépendance de la fréquence fondamentale avec la pression sous-glottique peut être modélisée par l'introduction d'une non-linéarité dans l'expression de la force élastique. Une étude analytique de Assaneo et Trevisan [AT13] explique que la fréquence fondamentale est toujours positivement corrélée à la pression sous-glottique lorsque que la force est exprimée à l'aide d'une non-linéarité cubique par rapport au déplacement, ce qui n'est pas le cas lorsque l'expression de la force est linéaire. Ce résultat est notamment confirmé par des simulations utilisant le modèle de Lous [LHVH98] : à partir d'un certain seuil, une augmentation de la pression sous-glottique provoque une diminution de la fréquence fondamentale [SD04].

Pour imiter des conditions de phonation représentatives du mécanisme I associé à une variation de $2,5 \cdot 10^{-2} \text{ Hz} \cdot \text{Pa}^{-1}$, Ishikaza et Flanagan [IF72] choisissent $\eta_k = 1 \cdot 10^6 \text{ m}^{-2}$.

2.3.5 Modélisation de la collision

La collision entre les deux cordes vocales opposées se produit au niveau de la masse i , à l'abscisse x_i lorsque $h_i \leq 0$. Cette condition peut également s'exprimer par l'inégalité suivante : $y_i \leq -y_0$. Le contact entre les deux cordes vocales est modélisé par une modification temporaire de leurs propriétés visqueuses et élastiques.

Force visqueuse

Dans le cas où la collision se produit au niveau de la masse i , la valeur de l'amortissement relatif ξ est remplacée par la valeur ξ_c . Le comportement du coefficient d'amortissement r_i s'écrit alors :

$$\begin{aligned} r_i &= 2 \xi \sqrt{k_i m_i} \text{ pour } y_i > -y_0, \\ r_i &= 2 \xi_c \sqrt{k_i m_i} \text{ pour } y_i \leq -y_0. \end{aligned} \quad (2.20)$$

Les valeurs de ξ_c utilisées dans la littérature sont légèrement supérieures à la valeur pour laquelle l'amortissement du système critique : $\xi_c = 1$.

Force élastique

Lors de la collision au niveau de la masse i , l'expression de la force élastique devient :

$$F_{ki}(y_i) = k_i y_i (1 + \eta_k y_i^2) + \alpha_{kc} k_i (y_i + y_0) (1 + \alpha_{\eta kc} \eta_k (y_i + y_0)^2). \quad (2.21)$$

La figure 2.4 présente la caractéristique de la force $F_{ki}(y_i)$. Cette expression de la force lors de la collision présente l'avantage d'assurer la continuité de la force élastique.

Limites

Cette modélisation de la collision a uniquement pour but d'augmenter la force de rappel en cas de collision, dans le but de limiter la durée de la phase pendant laquelle la glotte est fermée. Bien que les méthodes basées sur les éléments finis permettent de modéliser les interactions mécaniques liées au contact entre les cordes vocales [Gun03], ce type de modélisation n'a pas encore été proposé pour les modèles mécaniques simplifiés.

La géométrie bi-dimensionnelle du modèle à deux masses impose également une limite pour la modélisation de la collision. L'hypothèse selon laquelle la dimension antéro-postérieure de la

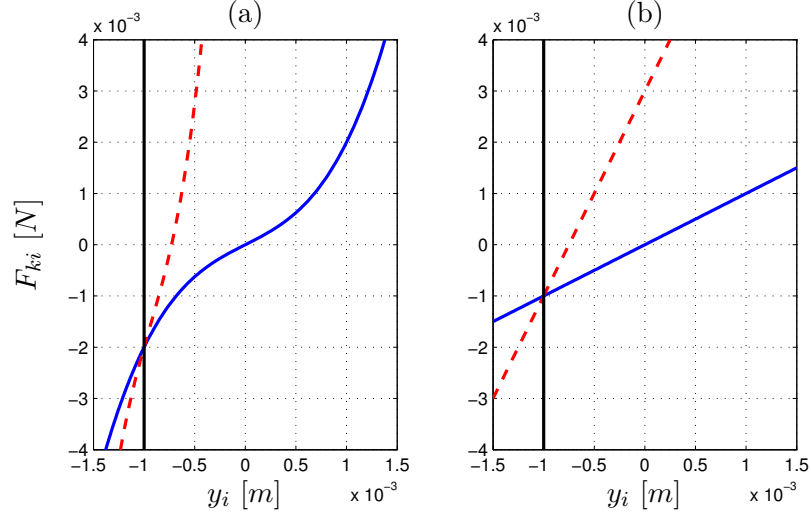


FIGURE 2.4 – Force élastique F_{ki} exercée par le ressort en fonction du déplacement de la masse m_i par rapport à sa position d'équilibre, pour $y_0 = 1$ mm, $\alpha_{kc} = 3$ et $\alpha_{\eta kc} = 5$. (a) : Cas non linéaire ($\eta_k = 10^6$ m $^{-2}$). (b) : Cas linéaire ($\eta_k = 0$ m $^{-2}$). La droite verticale symbolise la valeur du déplacement $y_i = -y_0$ pour lequel la collision se produit, la courbe en traits plein représente F_{ki} en absence de collision et la courbe en pointillés représente F_{ki} lorsque la collision a lieu.

constriction glottique est constante implique que l'indicateur de collision est binaire, la glotte est ouverte ou bien complètement fermée, alors que les observations *in vivo* [Cra87] indiquent que la collision se produit graduellement de l'extrémité des cordes vocales vers le centre.

2.4 Modèles acoustiques

La modélisation de la propagation des ondes acoustiques dans la trachée et le conduit vocal permet d'une part de tenir compte de l'effet des interactions acoustiques entre les cordes vocales et les résonateurs sur l'auto-oscillation (couplage acoustique) et d'autre part de déterminer la pression acoustique rayonnée au niveau des lèvres. La réalisation de simulations numériques temporelles nécessite de discrétiser le problème. La méthode utilisée ici est la simulation par guides d'onde numériques.

Simulation par guides d'onde numériques

La technique de simulation par guides d'onde numériques repose sur deux éléments :

- la discrétisation de la géométrie des résonateurs acoustiques par un réseau de tronçons à symétrie axiale (cylindres, cônes, ...), de longueur L constante,
- un changement de variable de pression et débit vers des variables de type ondes progressives dont la propagation est découplée dans les tronçons choisis.

La fréquence d'échantillonnage de la simulation est fixée de manière à ce que le pas de temps soit égal au temps de propagation des ondes acoustiques entre les deux extrémités d'un tronçon telle que $Fe = c/L$.

Il existe plusieurs types de discrétisation, associés à différents changements de variables, qui permettent de réaliser des simulations par guides d'onde numériques. Le principe de la simulation par ligne à retard, initialement utilisé en électronique, a été repris par Kelly et Lochbaum [KL62] pour synthétiser la voix humaine. Le conduit vocal est alors approximé par une agrégation de cylindres dans lesquels se propagent des ondes planes progressives. L'utilisation de cônes pour représenter la géométrie d'un résonateur acoustique a ensuite été introduite par Välimäki [Val95a]. Cette discrétisation est associée à l'utilisation d'ondes sphériques progressives et à un modèle de propagation basé sur l'équation des pavillons [Web19]. Enfin, Mignot [Mig09] utilise des ondes progressives dont la propagation est découplée dans les tronçons à courbure constante ($\partial_x^2 A(x) = cte$).

Ces différents degrés de raffinement dans l'approximation du rayon des résonateurs permettent d'améliorer le réalisme de la simulation tout en conservant un faible coût de calcul. Cependant, des problèmes d'instabilité apparaissent au niveau des fonctions de réflexion qui caractérisent les jonctions entre deux cônes décrivant une forme convexe [MA88, GKP90, Smi96, Ber99, MHM08] ou au niveau de la propagation dans un tube à courbure constante négative [Mig09]. Contrairement à la majorité des instruments à vent, pour lesquels l'évolution de la perce (de l'embouchure vers le pavillon) est toujours croissante, dans le cas de la parole les différentes constriction du conduit vocal impliquent que la condition de courbure positive n'est pas respectée. Cet obstacle motive ici le choix d'une discrétisation des résonateurs amont et aval par une agrégation de cylindres associée à l'utilisation des ondes plane progressives.

2.4.1 Modèle de propagation

Le modèle de propagation acoustique est un modèle mono-dimensionnel, reposant sur la propagation de fronts d'ondes perpendiculaires à l'axe principal du guide d'onde. Cette hypothèse d'ondes planes est valide pour des fréquences inférieures à la fréquence propre du premier mode propagatif non-plan [Bru98]. En considérant un guide d'onde de section circulaire, la condition de validité de cette hypothèse se traduit par :

$$f < \frac{\gamma_{10}c}{2\pi r}, \quad \gamma_{10} = 1,84. \quad (2.22)$$

La variable r est le rayon du guide d'onde et γ_{10} désigne le premier zéro de la dérivée de la fonction de Bessel de première espèce et d'ordre 1. En choisissant une valeur maximum de section égale à $A_{max} = 10^{-3} \text{ m}^2$, ce qui correspond au maximum de la section du conduit vocal lors de la production de la voyelle /a/ (au niveau de la cavité orale), l'équation donne une fréquence maximum égale à 5,7 kHz.

Cette hypothèse basse fréquence implique qu'il n'est pas nécessaire de reproduire les détails

de faible dimension de la géométrie du conduit vocal et de la trachée. De plus, le modèle de propagation étant mono-dimensionnel, la courbure du conduit vocal n'est pas considérée. Les autres hypothèses formulées sont que les parois du conduit vocal et de la trachée sont infiniment rigides et que les gradients de température sont négligeables dans l'ensemble de l'appareil phonatoire.

2.4.1.1 Équations de propagation

L'état acoustique à l'abscisse x est défini à l'aide des variables de pression $p(x, t)$ et de débit acoustique $u(x, t)$. Pour un instant donné, $p(x, t)$ et $u(x, t)$ sont uniformes sur l'ensemble de la section $A(x)$ du tube.

Le modèle de propagation acoustique consiste en l'association de l'équation d'onde et de l'équation d'Euler, linéarisées au premier ordre. Dans le domaine de Laplace, en supposant qu'à l'instant initial $p(x, 0) = u(x, 0) = 0$, le modèle s'écrit :

$$\left(\partial_x^2 - \left(\frac{s}{c} \right)^2 \right) p(x, s) = 0, \quad (2.23)$$

$$\frac{s\rho}{A(x)} u(x, s) + \partial_x p(x, s) = 0, \quad (2.24)$$

où s désigne la variable de Laplace. La résolution de l'équation 2.23 conduit à l'obtention d'une racine double en $\pm s/c$ pour l'expression de $p(x, s)$. L'expression de $u(x, s)$ est ensuite déduite de celle de $p(x, s)$. Sous forme matricielle, la solution du système s'écrit :

$$\begin{pmatrix} p(x, s) \\ u(x, s) \end{pmatrix} = \mathbf{G}(x, s) \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}, \quad \text{avec } \mathbf{G}(x, s) = \begin{pmatrix} e^{\frac{s}{c}x} & e^{-\frac{s}{c}x} \\ -\frac{1}{Z(x)} e^{\frac{s}{c}x} & \frac{1}{Z(x)} e^{-\frac{s}{c}x} \end{pmatrix}, \quad (2.25)$$

où C_1 et C_2 sont des constantes dépendantes des conditions initiales et $Z(x)$ est l'impédance acoustique telle que $Z(x) = \rho c / A(x)$.

2.4.1.2 Matrice de transport pour un cylindre

La matrice de transport $\mathbf{T}(L, s)$ traduit la relation entre les états acoustiques à chaque extrémité d'un tronçon, en $x = 0$ et $x = L$, telle que :

$$\begin{pmatrix} p_m(0, s) \\ u_m(0, s) \end{pmatrix} = \mathbf{T}(L, s) \begin{pmatrix} p_m(L, s) \\ u_m(L, s) \end{pmatrix}, \quad \text{avec } \mathbf{T}(L, s) = \begin{pmatrix} \cosh(sL/c) & Z(L) \sinh(sL/c) \\ \frac{1}{Z(0)} \sinh(sL/c) & \frac{Z(L)}{Z(0)} \cosh(sL/c) \end{pmatrix}. \quad (2.26)$$

L'expression de $\mathbf{T}(L, s)$ est calculée à partir de la matrice $\mathbf{G}(x, s)$ à l'aide de la relation $\mathbf{T}(L, s) = \mathbf{G}(0, s)\mathbf{G}^{-1}(L, s)$. Pour un tronçon cylindrique, la section $A(x)$ est constante, les impédances acoustiques à chaque extrémité sont donc également constantes et $Z(L) = Z(0)$.

2.4.1.3 Ondes planes progressives

Le changement de variables de pression et débit acoustique vers les ondes progressives est adapté à la simulation par guides d'ondes. Nous utilisons ici les ondes planes progressives notées p^+ et p^- , qui voyagent respectivement dans le sens des x croissants et décroissants. Ce changement de variables est donné par la matrice de passage $\mathbf{P}(x)$, définie par :

$$\begin{pmatrix} p^+(x, s) \\ p^-(x, s) \end{pmatrix} = \mathbf{P}(x) \begin{pmatrix} p(x, s) \\ u(x, s) \end{pmatrix}, \text{ avec } \mathbf{P}(x) = \frac{1}{2} \begin{pmatrix} 1 & Z(x) \\ 1 & -Z(x) \end{pmatrix}. \quad (2.27)$$

La matrice de transport $\mathbf{T}_p(L, s)$ donne la relation entre les états acoustiques décrits en fonction des ondes p^+ et p^- , à chaque extrémité du tronçon, telle que :

$$\begin{pmatrix} p^+(0, s) \\ p^-(0, s) \end{pmatrix} = \mathbf{T}_p(L, s) \begin{pmatrix} p^+(L, s) \\ p^-(L, s) \end{pmatrix}, \text{ avec } \mathbf{T}_p(L, s) = \begin{pmatrix} e^{sL/c} & 0 \\ 0 & e^{-sL/c} \end{pmatrix}. \quad (2.28)$$

L'expression de la matrice $\mathbf{T}_p(L, s)$ se calcule à partir de la matrice de passage \mathbf{P} et de la matrice de transport \mathbf{T} telle que $\mathbf{T}_p(L, s) = \mathbf{P}(0)\mathbf{T}(L, s)\mathbf{P}^{-1}(L)$. La version causale $\tilde{\mathbf{T}}_p(L, s)$ de la matrice $\mathbf{T}_p(L, s)$ donne l'expression des ondes progressives sortantes en fonction des ondes entrantes dans le tronçon telle que

$$\begin{pmatrix} p^+(L, s) \\ p^-(0, s) \end{pmatrix} = \tilde{\mathbf{T}}_p(L, s) \begin{pmatrix} p^+(0, s) \\ p^-(L, s) \end{pmatrix}, \text{ avec } \tilde{\mathbf{T}}_p(L, s) = \begin{pmatrix} e^{-sL/c} & 0 \\ 0 & e^{-sL/c} \end{pmatrix}. \quad (2.29)$$

La matrice $\tilde{\mathbf{T}}_p(L, s)$ est diagonale et les termes diagonaux sont des opérateurs qui représentent uniquement le retard L/c dû à la propagation des ondes d'une extrémité à l'autre du tronçon. La figure 2.5 illustre cette propagation.

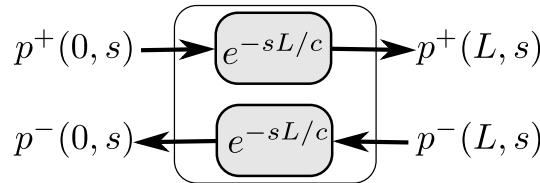


FIGURE 2.5 – Diagramme représentant la propagation découplée des ondes progressives planes p^\pm dans un cylindre.

2.4.1.4 Agrégation de tubes cylindriques

Les résonateurs situés respectivement en amont et en aval du larynx, sont modélisés par une agrégation de M_{am} et M_{av} tronçons cylindriques. Chaque tronçon est repéré par son indice m_{am} ou m_{av} (figure 2.6).

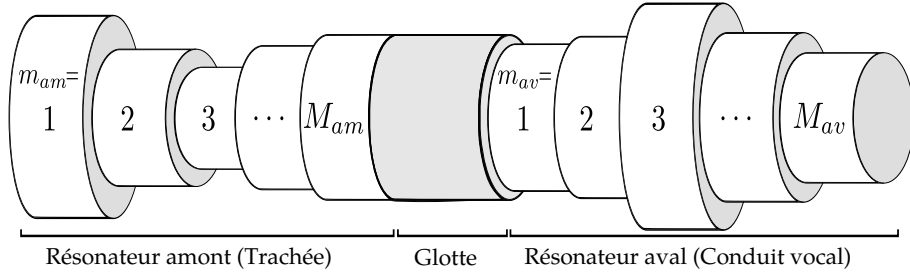


FIGURE 2.6 – Schéma représentant l'approximation de la trachée et du conduit vocal par une agrégation de tubes cylindriques.

2.4.1.5 Jonction entre deux tubes

Chaque tronçon d'indice m est caractérisé par sa section A_m et par son impédance acoustique $Z_m = \rho c / A_m$. À la jonction entre deux cylindres notés m et $m + 1$ (figure 2.7), la continuité des pressions et débits acoustiques entre l'extrémité droite du cylindre m (pour $x = L$) et l'extrémité gauche du cylindre $m + 1$ (pour $x = 0$), s'écrit :

$$\begin{pmatrix} p_m(L, s) \\ u_m(L, s) \end{pmatrix} = \begin{pmatrix} p_{m+1}(0, s) \\ u_{m+1}(0, s) \end{pmatrix}. \quad (2.30)$$

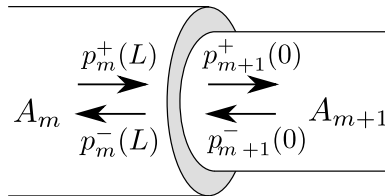


FIGURE 2.7 – Schéma représentant la jonction entre deux cylindres de section A_m et A_{m+1} .

La matrice \mathbf{J}_m définie par la relation $\mathbf{J}_m = \mathbf{P}_m \mathbf{P}_{m+1}^{-1}$ traduit la même condition, exprimée en fonction des ondes planes progressives telle que

$$\begin{pmatrix} p_m^+(L, s) \\ p_m^-(L, s) \end{pmatrix} = \mathbf{J}_m \begin{pmatrix} p_{m+1}^+(0, s) \\ p_{m+1}^-(0, s) \end{pmatrix}, \text{ avec } \mathbf{J}_m = \frac{1}{2Z_{m+1}} \begin{pmatrix} Z_{m+1} + Z_m & Z_{m+1} - Z_m \\ Z_{m+1} - Z_m & Z_{m+1} + Z_m \end{pmatrix}. \quad (2.31)$$

La version causale $\tilde{\mathbf{J}}_m$ de la matrice \mathbf{J}_m donne l'expression des ondes progressives réfléchies au niveau de la jonction en fonction des ondes incidentes :

$$\begin{pmatrix} p_{m+1}^+(0, s) \\ p_m^-(L, s) \end{pmatrix} = \tilde{\mathbf{J}}_m \begin{pmatrix} p_m^+(L, s) \\ p_{m+1}^-(0, s) \end{pmatrix}, \text{ avec } \tilde{\mathbf{J}}_m = \begin{pmatrix} 1 + k_m & -k_m \\ k_m & 1 - k_m \end{pmatrix}. \quad (2.32)$$

La matrice $\tilde{\mathbf{J}}_m$ fait apparaître une factorisation par la fonction de réflexion k_m , qui se présente comme un simple gain, fonction des sections A_m et A_{m+1} , telle que

$$k_m = \frac{Z_{m+1} - Z_m}{Z_{m+1} + Z_m} = \frac{A_m - A_{m+1}}{A_m + A_{m+1}}. \quad (2.33)$$

Cette factorisation, caractéristique d'une structure de Kelly-Lochbaum [KL62], permet de réduire le temps de calcul de la simulation. Cette structure est représentée en figure 2.8.

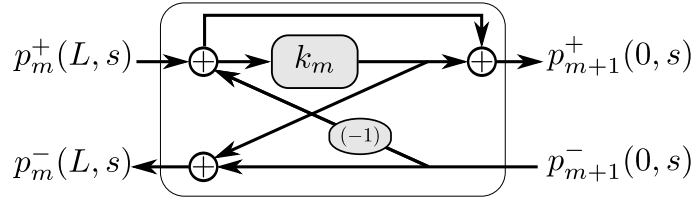


FIGURE 2.8 – Diagramme de la structure de Kelly-Lochbaum représentant la jonction entre deux cylindres.

2.4.2 Pertes viscothermiques

Le modèle de propagation acoustique présenté dans la partie précédente décrit un système acoustique conservatif. En réalité, la propagation acoustique dans un tuyau aux parois rigides est sujet à des phénomènes dissipatifs, en particulier aux effets de viscosité et de conduction thermique qui se produisent à proximité des parois. Chaigne et Kergomard [CK08] fournissent une formule approchée de la constante de propagation Γ pour une impulsion se propageant dans un cylindre, qui s'écrit :

$$\Gamma(s) = \frac{s}{c} + \frac{\kappa_0}{a} \sqrt{\frac{s}{c}}, \text{ avec } \kappa_0 = \sqrt{\ell_v} + (\gamma - 1) \sqrt{\ell_h}, \quad (2.34)$$

où ℓ_v et ℓ_h désignent respectivement les longueurs caractéristiques des effets visqueux et thermiques ($\ell_v = 4,0 \cdot 10^{-8} \text{ m}$, $\ell_h = 5,6 \cdot 10^{-8} \text{ m}$), γ le rapport des chaleurs spécifiques ($\gamma = 1,4$) et a le rayon du tronçon cylindrique tel que $a = \sqrt{A/\pi}$.

2.4.2.1 Validité du modèle

Cette approximation des effets viscothermiques sur la propagation des ondes planes repose sur l'hypothèse de tuyaux larges qui s'écrit :

$$\left| \sqrt{\frac{s}{c \ell_v}} a \right| \gg 1. \quad (2.35)$$

En choisissant $\left| \sqrt{\frac{s}{c \ell_v}} a \right| = 10$ et une fréquence minimum de validité du modèle de 80 Hz, l'approximation est valide pour des sections de tube supérieures à $A_{min} \approx 10^{-5} \text{ m}^2$. En présence des pertes viscothermiques, la matrice de transport \mathbf{T} définie par l'équation 2.26 devient pour un cylindre d'impédance caractéristique $Z = \rho c / (\pi a^2)$:

$$\mathbf{T}(L, s) = \begin{pmatrix} \cosh(\Gamma(s)L) & Z \sinh(\Gamma(s)L) \\ \frac{1}{Z} \sinh(\Gamma(s)L) & \cosh(\Gamma(s)L) \end{pmatrix}. \quad (2.36)$$

La matrice de transport exprimée en fonction des ondes p^\pm devient alors

$$\tilde{\mathbf{T}}_p(L, s) = \begin{pmatrix} e^{-\Gamma(s)L} & 0 \\ 0 & e^{-\Gamma(s)L} \end{pmatrix}. \quad (2.37)$$

La matrice $\tilde{\mathbf{T}}_p$ est diagonale ce qui signifie que la propagation des ondes p^\pm est découplée. Ce découplage est dû à la construction du modèle de pertes viscothermiques utilisé qui considère la propagation d'une impulsion seule (sans onde retour). Il est important de préciser que cette formulation n'est pas solution d'un système d'équations visant à modéliser la propagation acoustique et les pertes viscothermiques, tels que ceux proposés dans la littérature [Pol91, Mat94, H  l02, HHM03, HM06a]. Le terme $e^{-\Gamma(s)L}$ se d  compose en deux termes dont le premier repr  sente la propagation    travers le cylindre, sans modification de la forme d'onde, et le second l'effet induit par la viscosit   et l'induction thermique :

$$e^{-\Gamma(s)L} = e^{-\frac{s}{c}L} \check{D}(s), \text{ avec } \check{D}(s) = e^{-\epsilon \sqrt{s}}, \quad (2.38)$$

o   $\epsilon = -\frac{\kappa_0}{a\sqrt{c}}L$. La fonction de transfert $\check{D}(s)$ fait intervenir une puissance de la variable de Laplace s    exposant fractionnaire. Ce type de fonction de transfert appartient    la classe des fonctions de transfert irrationnelles. L'op  rateur de propagation $e^{-\frac{s}{c}L}$ est int  gr   au formalisme de la simulation par guides d'onde num  riques tandis que l'effet des pertes viscothermiques est mod  lis   par un filtre dont la fonction de transfert est $\check{D}(s)$ [Ned98, Ker91, Mat94, Duc01]. La figure 2.9 illustre la structure correspondante.

En posant $s = j\omega$ o   j d  signe le nombre imaginaire d  fini par $j = \sqrt{-1}$, ω la pulsation et en utilisant la relation $\sqrt{j} = (1 + j)/\sqrt{2}$, il est possible d'exprimer la fonction de transfert $\check{D}(j\omega)$ comme le produit d'un terme dissipatif et d'un terme dispersif telle que :

$$\check{D}(j\omega) = e^{-\epsilon\sqrt{j\omega}} = e^{-\epsilon\sqrt{\frac{\omega}{2}}} e^{-j\epsilon\sqrt{\frac{\omega}{2}}}, \text{ pour } \omega \geq 0. \quad (2.39)$$

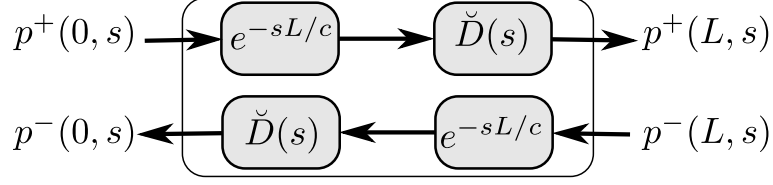


FIGURE 2.9 – Diagramme représentant la propagation découplée des ondes progressives planes p^\pm dans un cylindre.

2.4.2.2 Représentations diffusives

La construction d'un filtre numérique approximant les fonctions de transfert irrationnelles peut-être réalisée à l'aide des représentations diffusives (également appelées représentations intégrales) [HM06b, HM06a]. Le principe de ces représentations est d'exprimer un système complexe par une somme infinie de systèmes simples, du premier ordre, tel que :

$$\check{D}(s) = e^{-\epsilon\sqrt{s}} = s \int_0^{+\infty} \frac{\check{\mu}(\zeta)}{s + \zeta} d\zeta + \check{D}(0), \text{ avec } \check{\mu}(\zeta) = \frac{\sin(\epsilon\sqrt{\zeta})}{\pi\zeta} \text{ et } \check{D}(0) = 1. \quad (2.40)$$

Pour un nombre K fini de systèmes, la somme s'écrit :

$$\tilde{D}(s) = s \sum_{k=1}^K \frac{\check{\mu}_k}{s + \zeta_k} + \check{D}(0) = - \sum_{k=1}^K \frac{\zeta_k \check{\mu}_k}{s + \zeta_k} + \left(\sum_{k=1}^K \check{\mu}_k + \check{D}(0) \right). \quad (2.41)$$

Les paramètres ζ_k et μ_k sont deux réels qui représentent respectivement les pôles et les poids accordés à chaque système du premier ordre.

2.4.2.3 Simulation temporelle

Dans le domaine temporel, le système d'entrée \check{u} et de sortie \check{y} est décrit par la fonction de transfert $\check{D}(s)$ dont l'état est décrit par les sous-états ψ_k avec $1 \leq k \leq K$, tel que :

$$\partial_t \psi_k(t) = -\zeta_k \psi_k(t) + \check{u}(t), \quad 1 \leq k \leq K, \quad (2.42)$$

$$\check{y}(t) = \sum_{k=1}^K (-\zeta_k \check{\mu}_k) \psi_k(t) + \left(\check{D}(0) + \sum_{k=1}^K \check{\mu}_k \right) \check{u}(t). \quad (2.43)$$

En prenant $\check{u}[n] \approx \check{u}(nT_e)$ où T_e est la période d'échantillonnage, le système discrétisé s'écrit :

$$\psi_k[n] = e^{-\zeta_k T_e} \psi_k[n-1] + \frac{e^{-\zeta_k T_e} - 1}{-\zeta_k} \check{u}[n-1], \quad 1 \leq k \leq K, \quad (2.44)$$

$$\check{y}[n] = \sum_{k=1}^K (-\zeta_k \check{\mu}_k) \psi_k[n] + \left(\check{D}(0) + \sum_{k=1}^K \check{\mu}_k \right) \check{u}[n]. \quad (2.45)$$

2.4.2.4 Optimisation de l'approximation

Bien que le poids $\check{\mu}(\zeta_k)$ puisse être calculé analytiquement par une intégrale dans le plan complexe, Hélie et Matignon [HM06a] utilisent une méthode alternative qui donne de meilleures approximations pour de faibles valeurs de K . La méthode d'optimisation proposée est basée sur l'algorithme des moindres carrés. Le critère $C(\boldsymbol{\mu})$ à minimiser est calculé pour un nombre fini de pulsation ω_n avec $1 \leq n \leq N$. Le critère $C(\boldsymbol{\mu})$ est défini comme la distance entre la fonction de transfert $\check{D}(j\omega_n)$ et son approximation $\check{D}_{\boldsymbol{\mu}}(j\omega_n)$, pondérée par le poids $w_D(\omega_n)$ tel que :

$$C(\boldsymbol{\mu}) = \sum_{n=1}^N |(\check{D}_{\boldsymbol{\mu}}(j\omega_n) - \check{D}(j\omega_n)) w_D(\omega_n)|^2 [\ln \omega_{n+1} - \ln \omega_n], \text{ avec } w_D(\omega_n) = \frac{\omega_n}{1 + \omega_n}. \quad (2.46)$$

Sous forme matricielle, $C(\boldsymbol{\mu})$ s'écrit :

$$C(\boldsymbol{\mu}) = (\mathbf{M}\boldsymbol{\mu} - \mathbf{d})^* \mathbf{W}^* \mathbf{W}(\mathbf{M}\boldsymbol{\mu} - \mathbf{U}), \quad (2.47)$$

où la notation \mathbf{M}^* désigne la matrice conjuguée transposée telle que $\mathbf{M}^* = \overline{\mathbf{M}}^T$. Les différentes matrices et vecteurs utilisés sont définis dans le tableau 2.2.

Éléments	Dimensions	Caractéristiques	Définition
\mathbf{M}	$N \times K$	matrice imaginaire	$\mathbf{M}[n, k] = (j\omega_n + \zeta_k)^{-1}$
\mathbf{W}	$N \times N$	matrice réelle diagonale	$\mathbf{W}[n, n] = w_D(\omega_n) \sqrt{\ln \omega_{n+1} - \ln \omega_n}$
\mathbf{d}	$N \times 1$	vecteur complexe	$\mathbf{d}[n] = \check{D}(j\omega_n)$
$\boldsymbol{\mu}$	$1 \times K$	vecteur réel	$\boldsymbol{\mu}[k] = \check{\mu}_k$

TABLE 2.2 – Récapitulatif des éléments vectoriels utilisés pour l'optimisation de l'approximation par la méthode des moindres carrés.

En imposant que le vecteur $\boldsymbol{\mu}$ soit réel, la résolution du problème aux moindres carrés s'écrit :

$$\mu = [\Re(\mathbf{M}^* \mathbf{W}^* \mathbf{M} \mathbf{W})]^{-1} + \Re(\mathbf{M}^* \mathbf{W}^* \mathbf{M} \mathbf{d}). \quad (2.48)$$

Les N valeurs de ω_n et les K valeurs de ζ_k sont réparties de façon logarithmique entre les pulsations $\omega_1 = 2\pi \text{ rad.s}^{-1}$ et $\omega_{K+1} = \pi F_e \text{ rad.s}^{-1}$. Les valeurs des paramètres utilisés pour l'approximation de la fonction de transfert $\check{D}(s)$ sont $K = 4$ et $N = 32$. La figure 2.10 illustre l'approximation de la fonction de transfert \check{D} pour les valeurs des sections maximum et minimum, imposées respectivement par l'approximation de propagation en onde plane (équation 2.22) et par l'approximation de large tuyaux pour le modèle de pertes viscothermiques (équation 2.35).

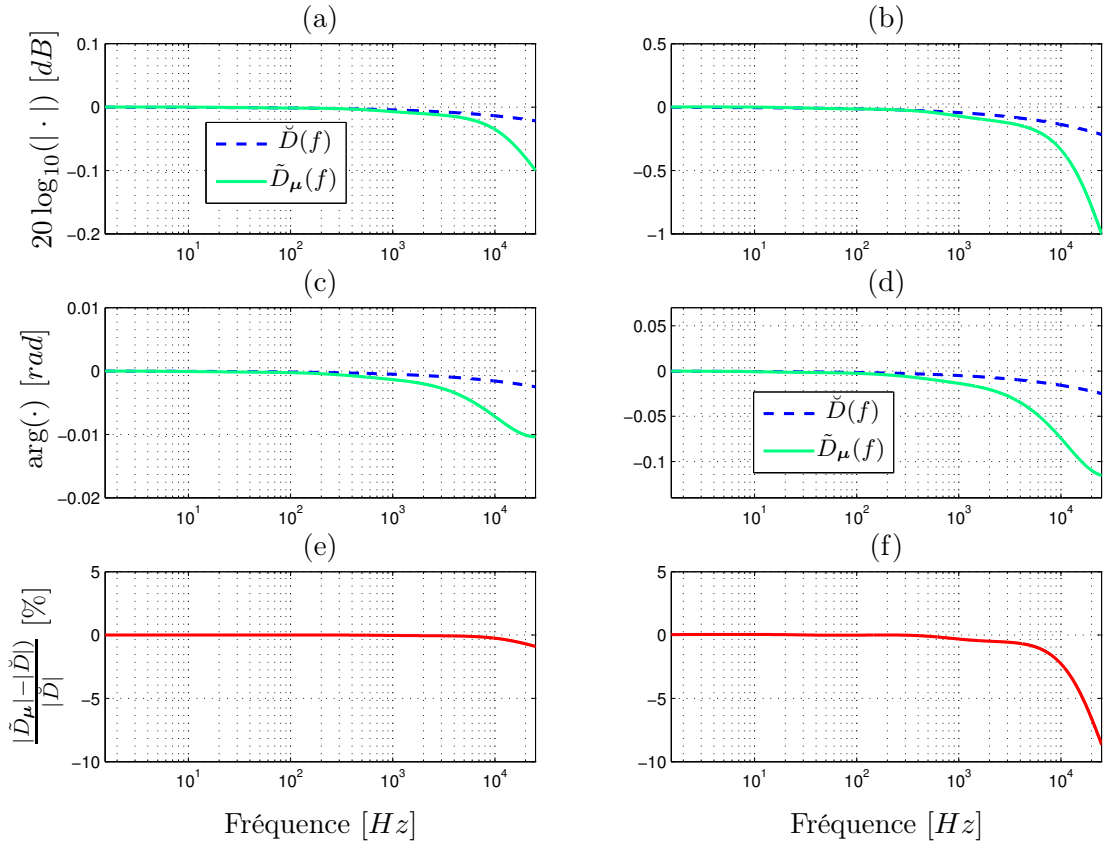


FIGURE 2.10 – Comparaison dans le domaine fréquentielle de $\check{D}(f)$ et de son approximation $\tilde{D}_\mu(f)$ pour un cylindre de longueur $L = 1 \text{ cm}$, de section $A_{max} = 10^{-3} \text{ m}^2$ (colonne de gauche) et $A_{min} = 10^{-5} \text{ m}^2$ (colonne de droite), en prenant $F_e = 50 \text{ kHz}$, $K = 4$ et $N = 32$. Modules (a,b) et phases (c,d) des fonctions de transfert. Erreur relative d'approximation sur le module (e,f).

Pour des cylindres dont la section est inférieure à $A_{min} = 10^{-5} \text{ m}^2$, le modèle de pertes viscothermiques sur-estime les effets dissipatifs. En pratique, les effets viscothermiques lors de la propagation sont limités à ceux déterminés pour une section égale à A_{min} . Les effets dus aux pertes viscothermiques se traduisent essentiellement par une atténuation de

l'énergie contenue en haute-fréquence. Leur influence sur les fréquences proches de la fréquence fondamentale de la phonation, et donc sur l'auto-oscillation des cordes vocales, est donc faible. L'intérêt de modéliser les pertes viscothermiques est essentiellement d'améliorer le réalisme des simulations d'un point de vue perceptif.

2.4.3 Modèle de rayonnement

2.4.3.1 Impédance de rayonnement

À l'extrémité du conduit vocal, une partie de l'énergie acoustique est dissipée par le rayonnement des ondes acoustiques vers le milieu extérieur. Ce phénomène est modélisé par l'impédance de rayonnement. Une formulation analytique qui prendrait en compte la complexité de la géométrie du visage humain n'est pas envisageable. Pour modéliser la production de parole, le modèle de rayonnement le plus fréquemment utilisé est le piston circulaire encastré dans un écran de dimension infinie [Fla72, IF72, Sto95]. Cette simplification dans la modélisation de la géométrie du visage et de la tête humaine peut entraîner un décalage des fréquences de résonance du conduit vocal vers des valeurs plus élevées. Ce phénomène est d'autant plus important que la fréquence est élevée. Des simulations en trois dimensions utilisant les éléments finis montrent que, pour des fréquences inférieures à 10 kHz , ce décalage ne dépasse pas les 5 % [AGA13]. Dans le cas présent, il est essentiel que le modèle de rayonnement utilisé soit applicable à la simulation dans le domaine temporel.

L'impédance de rayonnement $Z_r(s)$ fait le lien entre la pression rayonnée au niveau des lèvres $p_M(L, s) = p_M^+(L, s) + p_M^-(L, s)$ et le débit acoustique à l'extrémité du résonateur $u_M(L, s)$, telle que $p_M(L, s) = Z_r(s)u_M(L, s)$. Pour les ondes p^\pm , l'impédance de rayonnement se traduit par une fonction de réflexion $R(s)$ et une fonction de transmission $T(s)$, définies par :

$$R(s) = \frac{p_M^-(L, s)}{p_M^+(L, s)} = \frac{Z_r(s) - Z_M}{Z_r(s) + Z_M}, \quad (2.49)$$

$$T(s) = \frac{p_M(L, s)}{p_M^+(L, s)} = \frac{2Z_r(s)}{Z_r(s) + Z_M} = R(s) + 1, \quad (2.50)$$

avec $Z_M = (\rho c)/A_M$.

La figure 2.11 illustre la structure utilisée pour la simulation.

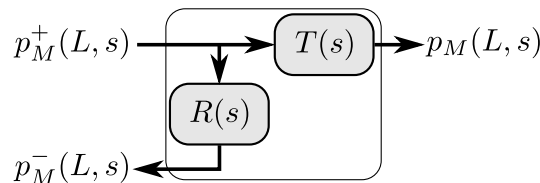


FIGURE 2.11 – Diagramme représentant la réflexion et la transmission des ondes progressives planes p^\pm à l'extrémité du résonateur.

2.4.3.2 Expression de la fonction de réflexion

L'expression utilisée pour la fonction de réflexion est la formule approchée proposée par Silva et al. [SGK⁺09]. En se basant sur les travaux de Zorumski [Zor73] et Norris et Sheng [NS89], les auteurs déterminent une approximation basse fréquence de la fonction de réflexion $R(j\omega)$ associée à l'impédance d'un piston plan, en imposant les propriétés de causalité et de symétrie hermitienne. La transformée de fourrier inverse de $R(j\omega)$ s'écrit :

$$r(t) = \begin{cases} \frac{c}{ad_2(\gamma_1 - \gamma_2)} \left((n_1\gamma_2 - 1)e^{-\gamma_2 \frac{ct}{a}} - (n_1\gamma_1 - 1)e^{\gamma_1 \frac{ct}{a}} \right) & \text{pour } t > 0, \\ 0 & \text{pour } t \leq 0, \end{cases} \quad (2.51)$$

avec $\gamma_{1,2} = \frac{d_1 \pm \sqrt{d_1^2 - 4d_2}}{2d_2}$, $n_1 = 0,182$, $d_1 = 1,825$ et $d_2 = 0,649$.

L'erreur relative d'approximation en utilisant cette formule reste faible (de l'ordre de 5 %) tant que $\omega a/c \leq 1$. La validité de cette approximation est limitée aux tuyaux de grande section. En prenant $A_{max} = 10^{-3} \text{ m}^2$, la formule approchée est valide pour des fréquences inférieures à 3 kHz.

2.4.3.3 Approximation de la fonction de réflexion

L'expression de la fonction de réflexion (équation 2.51) est utilisée pour concevoir un filtre numérique à l'aide de la méthode de l'invariance de la réponse impulsionnelle, telle que

$$\tilde{r}(t) = r(t) \sum_{p=0}^P \delta(t - pT_e). \quad (2.52)$$

Les coefficient $\tilde{r}[n]$ du filtre d'ordre P correspondant sont donnés par l'expression suivante :

$$\tilde{r}[n] = \frac{r(nT_e)}{\sum_{n=0}^P |r(nT_e)|}, \quad n \in \mathbb{N}, \quad 0 \leq n \leq P. \quad (2.53)$$

Le choix de l'ordre du filtre à réponse impulsionnelle finie repose sur un compromis entre la précision de l'approximation et le temps de calcul. Dans ces travaux, ce paramètre est fixé à $P = 12$. L'approximation ainsi réalisée est illustrée par la figure 2.12. L'erreur relative d'approximation sur le module et la phase augmente avec la fréquence. À la limite de validité de ce modèle de rayonnement ($f < 3 \text{ kHz}$), l'erreur d'approximation commise sur le module ne dépasse pas 7% dans le cas d'une section de $A_{max} = 10^{-3} \text{ m}^2$ et 0,5% dans le cas d'une section de $A_{min} = 10^{-5} \text{ m}^2$.

Bien que modéliser la fonction de réflexion par un filtre autorégressif permettrait de limiter considérablement l'ordre du filtre pour une précision d'approximation équivalente, l'utilisation d'un filtre à réponse impulsionnelle finie présente l'avantage de garantir la stabilité du système.

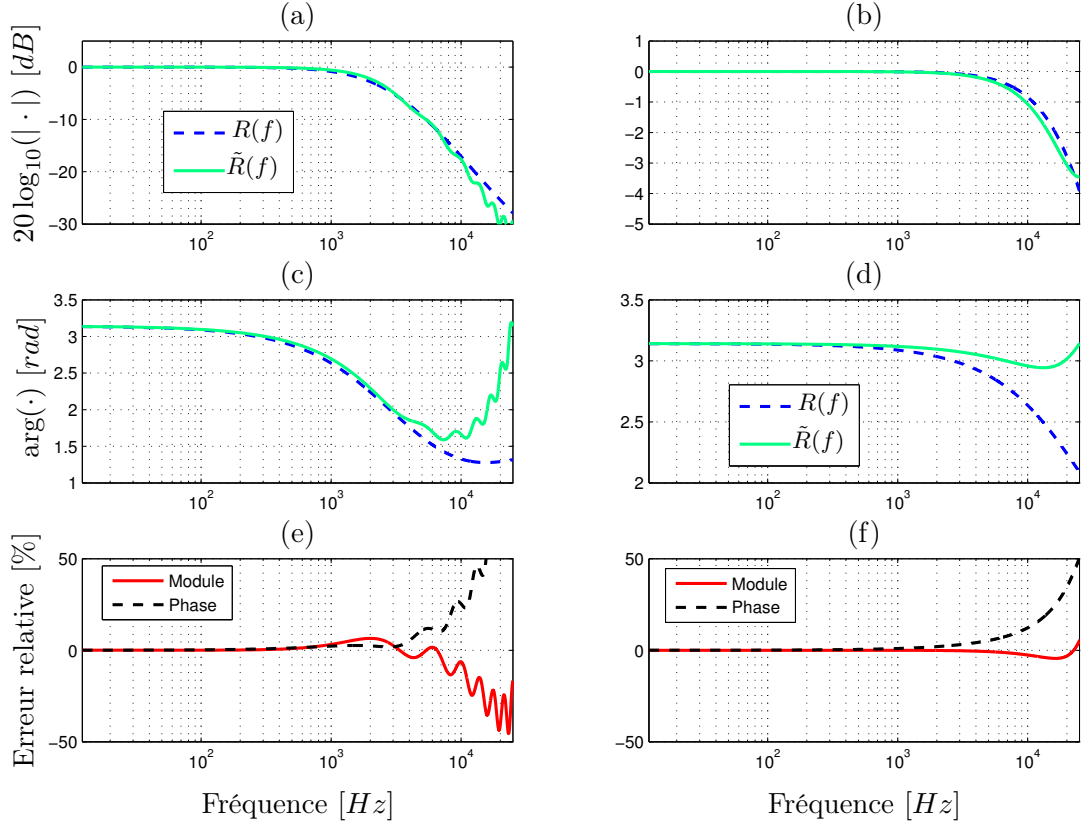


FIGURE 2.12 – Comparaison dans le domaine fréquentiel de la fonction de réflexion $R(f)$ et de son approximation $\tilde{R}(f)$ pour un cylindre de section $A_{max} = 10^{-3} \text{ m}^2$ (colonne de gauche) et $A_{min} = 10^{-5} \text{ m}^2$ (colonne de droite), encastré dans un écran infini avec $F_e = 50 \text{ kHz}$. Modules (a,b) et phases (c,d) des fonctions de transfert. Erreur relative d'approximation sur le module et la phase (e,f).

De plus, la méthode de l'invariance de la réponse impulsionnelle permet de limiter le nombre de calculs nécessaires à la détermination des coefficients $\tilde{r}[n]$, ce qui réduit considérablement de temps de calcul lorsque l'aire intéro-labiale varie au cours de la simulation, comme dans le cas des simulations de plosives bilabiales réalisées au chapitre 7.

2.4.4 Impédance d'entrée des résonateurs acoustiques

Le comportement acoustique d'un résonateur peut être caractérisé par son impédance d'entrée Z_e définie comme le rapport entre la pression et le débit acoustique à l'entrée du résonateur. L'impédance d'entrée d'un résonateur acoustique discrétisé peut-être déterminée analytiquement. La condition de continuité du débit et de la pression à la jonction entre chaque tronçon (équation 2.30) implique que Z_e se déduit du produit des matrices de transport \mathbf{T}_m (équation 2.36) et de l'impédance de rayonnement Z_r . À chaque extrémité d'un résonateur

composé de M tronçons cylindriques, pression et débit acoustique sont reliés par la relation :

$$\begin{pmatrix} p_1(0, s) \\ u_1(0, s) \end{pmatrix} = \mathbb{T}_M(s) \begin{pmatrix} p_M(L, s) \\ u_M(L, s) \end{pmatrix}, \text{ avec } \mathbb{T}_M = \prod_{m=1}^M \mathbf{T}_m(L, s). \quad (2.54)$$

Le rapport $p_1(0, s)/u_1(0, s)$ peut s'exprimer à partir des termes de la matrice $\mathbb{T}_M(s)$ (équation 2.54), l'impédance d'entrée s'écrit alors :

$$Z_e(s) = \frac{\mathbb{T}_M[1, 1]Z_r(s) + \mathbb{T}_M[1, 2]}{\mathbb{T}_M[2, 1]Z_r(s) + \mathbb{T}_M[2, 2]}, \quad (2.55)$$

où la notation $\mathbf{T}_p[i, j]$ désigne l'élément situé à la ligne i et à la colonne j de la matrice \mathbf{T}_p .

2.4.5 Couplage acoustique-aérodynamique

Le modèle aérodynamique qui suppose un écoulement incompressible et le modèle acoustique qui décrit la propagation d'une onde de compression, interagissent uniquement au niveau du larynx, de part et d'autre de la constriction glottique. Le couplage acoustique est basé sur la prise en compte des variations de pression acoustique en amont p_{am} et en aval p_{av} de la glotte pour le calcul du débit glottique U_g . Les pressions p_{am} et p_{av} correspondent aux pressions acoustiques aux extrémités des résonateurs amont et aval telles que $p_{am}(t) = p_{M_{am}}(L, t)$ et $p_{av}(t) = p_1(0, t)$.

À partir de l'expression de la matrice \mathbf{P}_m (équation 2.27) pour $m_{av} = 1$ et $m_{am} = M_{am}$, il est possible d'exprimer l'onde progressive réfléchie et la pression acoustique totale en fonction de l'onde incidente et du débit, en amont du larynx d'après :

$$\begin{pmatrix} p_{M_{am}}^-(L, t) \\ p_{am}(t) \end{pmatrix} = \begin{pmatrix} 1 & -Z_{M_{am}} \\ 2 & -Z_{M_{am}} \end{pmatrix} \begin{pmatrix} p_{M_{am}}^+(L, t) \\ U_g(t) \end{pmatrix} \quad (2.56)$$

et en aval du larynx d'après :

$$\begin{pmatrix} p_1^+(0, t) \\ p_{av}(t) \end{pmatrix} = \begin{pmatrix} 1 & Z_1 \\ 2 & Z_1 \end{pmatrix} \begin{pmatrix} p_1^-(0, t) \\ U_g(t) \end{pmatrix}. \quad (2.57)$$

En prenant en compte la contribution des pressions acoustiques, l'équation de Bernoulli corrigée (équation 2.15), appliquée entre les points d'abscisse x_0 et x_3 , devient :

$$P_{sub}(t) + p_{am}(t) = p_{av}(t) + \frac{\rho}{2} \left(\frac{U_g(t)}{L_g} \right)^2 \left(\frac{1}{h^2(x_s, t)} - \frac{1}{h_0^2(t)} \right) + \Delta P_\nu(x_s, t). \quad (2.58)$$

En utilisant les expressions de p_{am} et p_{av} données par les équations 2.56 et 2.57, l'équation 2.58 devient :

$$P_{sub}(t) + 2p_{Mam}^+(0, t) - Z_{Mam} U_g(t) = 2p_1^-(0, t) + Z_1 U_g(t) + \frac{\rho}{2} \left(\frac{U_g(t)}{L_g} \right)^2 \left(\frac{1}{h^2(x_s, t)} - \frac{1}{h_0^2(t)} \right) + \Delta P_\nu(x_s, t). \quad (2.59)$$

Le débit glottique U_g est déterminé en résolvant cette équation du second degré.

2.5 Implémentation numérique

L'algorithme utilisé pour réaliser des simulations temporelles discrètes de voyelles tenues est schématisé en figure 2.13.

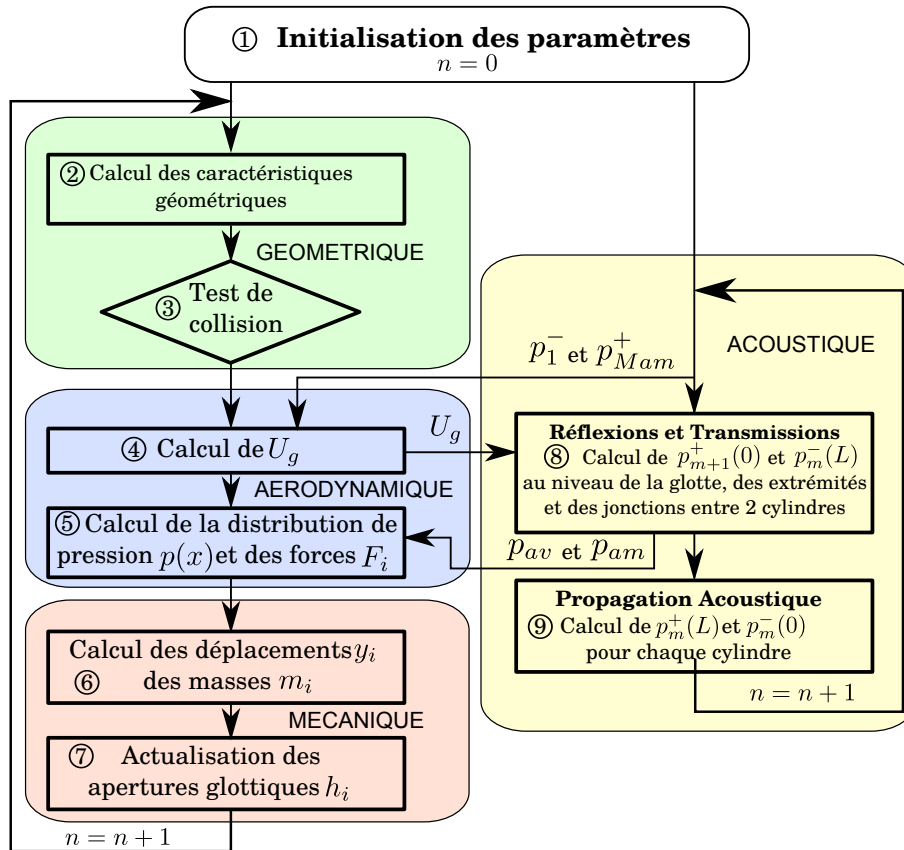


FIGURE 2.13 – Diagramme représentant l'algorithme de résolution temporelle du modèle à deux masses couplé avec les résonateurs acoustiques aval et amont, dans le cas d'une voyelle tenue et sans variation des paramètres d'entrée. La notation n désigne l'indice temporel discret.

Cette représentation met en avant le fait que la structure du code de simulation est compo-

sée de deux branches presque indépendantes. La première est constituée de la partie aérodynamique et mécanique du modèle à deux masses et la deuxième de la propagation acoustique dans les résonateurs amont et aval. Les interactions entre les parties acoustique et aérodynamique sont représentées sur le diagramme par des flèches horizontales.

Le fonctionnement de la branche aérodynamique et mécanique repose sur l'itération de deux étapes : pour une géométrie donnée, l'écoulement est déterminé à partir de la pression sous-glottique P_{sub} puis le déplacement des masses est calculé à partir des forces de pression aérodynamiques.

Les différentes étapes représentées par des blocs sur la figure 2.13 sont détaillées ci-dessous.

- ① L'initialisation inclut le choix des valeurs de pression P_{sub} , des paramètres géométriques et mécaniques du modèle à deux masses. Pour la partie acoustique, cette étape implique les calculs de filtres modélisant la réflexion aux extrémités des résonateurs et les pertes viscothermiques subies lors de la propagation dans chaque tronçon cylindrique.
- ② Le calcul des caractéristiques géométriques comprend le calcul de la position du point de séparation de l'écoulement glottique (abscisse x_s et aperture h_s), ainsi que le calcul de données sur la géométrie de la glotte, utiles pour déterminer le débit U_g et la distribution de pression $P(x)$.
- ③ Dans le cas où la collision a lieu, lorsque h_1 ou h_2 est inférieure ou égale à zéro, le comportement mécanique des cordes vocales (section 2.3.5) et la distribution de pression $P(x)$ sont modifiés. Le débit U_g est assigné à la valeur nulle.
- ④ Le débit U_g est calculé à partir de P_{sub} , de la géométrie glottique et des ondes progressives incidentes p_1^- et $p_{M_{am}}^+$ en résolvant l'équation 2.59.
- ⑤ La distribution de pression $P(x)$ dans la glotte est déterminée à partir de la géométrie, du débit glottique et des pressions acoustiques p_{am} et p_{av} de part et d'autre de la glotte. Les forces de pression F_i s'exerçant sur les parois de la glotte, peuvent ensuite être calculées à partir de l'équation 2.17. En pratique, la distribution $P(x)$ est déterminée sur 200 points et les forces de pression sont calculées par intégration numérique en utilisant la méthode des trapèzes.
- ⑥ Les déplacements y_1 et y_2 des masses m_1 et m_2 sont calculés à l'aide de l'équation 2.16 en utilisant un schéma de discrétisation de Runge-Kutta d'ordre 4.
- ⑦ En vue de l'itération suivante, les apertures h_1 et h_2 sont actualisées à l'aide des déplacements précédemment calculés.
- ⑧ Le calcul des réflexions et transmissions acoustiques correspond au calcul des ondes progressives $p_{m+1}^+(0)$ et $p_m^-(L)$ à partir des ondes $p_m^+(L, s)$ et $p_{m+1}^-(0, s)$ en utilisant l'équation 2.32. Cette étape inclut également le calcul des ondes réfléchies aux extrémités des résonateurs en utilisant le filtre numérique présenté en section 2.4.3 et le calcul des ondes $p_1^+(0)$ et $p_{M_{am}}^-(L)$ réfléchies au niveau de la glotte, en utilisant les matrices présentées par les équations 2.56 et 2.57. Ces mêmes matrices permettent également de calculer les pressions acoustiques p_{am} et p_{av} , requises pour l'étape ⑤.
- ⑨ L'étape de propagation consiste à calculer les ondes planes $p_m^+(L)$ et $p_m^-(0)$ pour l'itération suivante en fonction des ondes $p_m^+(0)$ et $p_m^-(L)$. Ce calcul fait intervenir le filtre numérique approximant l'effet des pertes viscothermiques, présenté en section 2.4.2.

2.6 Bilan

Bilan aérodynamique

La géométrie de la glotte est représentée par une constriction bi-dimensionnelle, de largeur constante, et dont la hauteur varie linéairement par morceaux dans la direction principale de l'écoulement. L'écoulement glottique, considéré comme incompressible, est décrit à l'aide d'un modèle quasi-stationnaire et unidimensionnel. Le modèle utilisé, basé sur l'équation de Bernoulli pour un fluide parfait, tient compte de la formation d'un jet en aval de la constriction. Les pertes visqueuses sont modélisées en utilisant l'expression du gradient de pression pour un écoulement de Poiseuille.

Bilan mécanique

Le comportement mécanique des cordes vocales est décrit à l'aide d'un modèle simplifié constitué de deux masses. En considérant l'approximation des éléments localisés, cette représentation est valable pour des fréquences inférieures à 10 kHz . Le modèle mécanique néglige l'effet des sollicitations passées, subies par le matériau. Les forces visqueuses et élastiques sont exprimées à l'aide de non-linéarités cubiques. La collision est modélisée par une modification temporaire des propriétés mécaniques des cordes vocales.

Bilan acoustique

Le modèle de propagation acoustique utilisé est mono-dimensionnel et basé sur une approximation basse fréquence. Le modèle est implémenté au moyen de guides d'onde numériques, basé sur la propagation d'ondes planes progressives et sur la discrétisation de la géométrie des résonateurs acoustiques. Les effets de la réflexion à l'extrémité des résonateurs acoustiques et des pertes viscothermiques lors de la propagation acoustique sont modélisés par des filtres numériques.

Pour des résonateurs dont les sections sont comprises entre $A_{min} = 10^{-5} \text{ m}^2$ et $A_{max} = 10^{-3} \text{ m}^2$, les modèles de propagation et de rayonnement sont valables pour des fréquences inférieures à 3 kHz et le modèle de pertes viscothermiques est valide pour des fréquences supérieures à 80 Hz . Les erreurs d'approximation induites par les filtres numériques restent faibles pour des fréquences inférieures à 1 kHz .

Contributions

Par rapport aux précédents travaux de thèse menés au département Parole et Cognition du Gipsa-Lab ¹ [Vil02, Rut07], les apports de ce travail de thèse au niveau des modèles utilisés pour les simulations numériques temporelles, résident dans les modélisations mécanique et acoustique.

Afin de mieux représenter le comportement mécanique des cordes vocales lors d'oscillations

1. anciennement l'Institut de la Communication Parlée

de forte amplitude, des termes non linéaires ont été ajoutés dans les expressions des forces visqueuses et élastiques. Pour la résolution numérique, l'introduction de ces non-linéarités mécaniques nécessite l'utilisation d'un schéma numérique explicite. Les déplacements sont donc calculés à l'aide d'un schéma de Runge-Kutta d'ordre 4, qui présente une stabilité plus importante qu'un schéma d'Euler explicite.

Au niveau de la modélisation acoustique, l'utilisation de la formule de Silva et al. [SGK⁺09] pour la conception d'un filtre numérique modélisant la réflexion aux extrémités des résonateurs, permet d'obtenir une réponse impulsionnelle réelle et causale. La garantie de la causalité du filtre constitue un avantage majeur par rapport à des méthodes basées sur l'utilisation d'une expression de la fonction de réflexion issue de l'approximation basse-fréquence des fonctions de Bessel et Struve. La dernière contribution apportée par ces travaux est la prise en compte de pertes viscothermiques au niveau des parois des résonateurs acoustique. Le formalisme des représentations diffusives permet d'approximer efficacement l'effet des pertes viscothermiques lors de la propagation, tout en limitant le coût de calcul [HM06b, HM06a].

Deuxième partie

Sauts de fréquence fondamentale et transitions de mécanisme laryngé

Description de la deuxième partie

Cette partie traite de la modélisation physique des transitions de mécanismes laryngés et des sauts de fréquence fondamentale qui les accompagnent.

L'objectif du chapitre 3 est d'observer expérimentalement un exemple de saut de fréquence produit involontairement. Dans cette optique, des mesures acoustique et électroglottographique sont réalisées sur un sujet sourd profond lors de la production d'une voyelle tenue.

Ensuite, les travaux présentés dans le chapitre 4 ont pour but de vérifier que les modèles physiques proposés dans le chapitre 2 sont aptes à reproduire ce phénomène de transition entre mécanismes ainsi que le saut de fréquence fondamentale associé. Dans un premier temps, des sauts de fréquence sont observés sur une maquette auto-oscillante de l'appareil phonatoire et dans un second temps, des simulations numériques tentant de reproduire ces observations sont réalisées.

Enfin, les simulations numériques sont utilisées dans le chapitre 5 afin d'étudier les stratégies motrices qui induisent les transitions de mécanismes laryngés et l'influence des différentes stratégies sur le saut de fréquence fondamentale.

À l'issue de ces trois chapitres, plusieurs hypothèses sur la stratégie employée par le sujet sourd lors de la production du saut de fréquence observé, sont proposées.

Bibliographie et observations *in vivo*

Sommaire

3.1	Introduction	57
3.2	Éléments bibliographiques	58
3.2.1	Observations de transitions de mécanisme laryngé	58
3.2.2	Hypothèses sur l'origine des transitions de mécanisme	59
3.2.3	Simulations numériques de bifurcations et de transitions de mécanisme laryngé	60
3.3	Observation <i>in vivo</i> de saut de fréquence fondamentale	62
3.3.1	Matériel et méthode	62
3.3.2	Résultats	64
3.3.3	Discussion et bilan	67

3.1 Introduction

Les transitions de mécanisme laryngé peuvent entraîner des variations importantes de la fréquence fondamentale de la voix. Lorsqu'elle sont produites de façon involontaire, ces variations excessives altèrent considérablement la communication. L'objectif de cette partie est de déterminer les stratégies motrices et les origines physiques de ce type de transition. L'approche abordée dans cette partie est d'observer le phénomène sur un locuteur sourd profond, ne disposant d'aucune perception auditive de ses propres productions vocales.

Étant donné la grande variabilité inter-individuelle des productions des locuteurs sourds et malentendants, la démarche employée consiste à disposer d'un exemple de mesure de sauts de fréquences produit involontairement dans un contexte phonétique simple. Cet exemple sera utilisé comme base de réflexion dans la suite de la partie.

La première section de ce chapitre est consacrée à une synthèse bibliographique des études théoriques, expérimentales et numériques sur les transitions de mécanisme laryngé. La deuxième section présente les mesures effectuées sur un sujet sourd profond post-linguistique.

3.2 Éléments bibliographiques

3.2.1 Observations de transitions de mécanisme laryngé

Les études réalisées sur des larynx vivants [SSM99, Hen01, Lam10] ou sur des larynx excisés [BHTS96, SSM99, AFS09] montrent que, dans de très nombreux cas, les transitions de mécanismes (1.1.4) sont accompagnées d'un saut de fréquence fondamentale. Le saut de fréquence est généralement positif lors d'une transition dans le sens MI→MII et négatif lors d'une transition dans le sens inverse (MII→MI).

Dans le cas de mesures réalisées sur des chanteurs lyriques, l'utilisation de l'EKG a permis de mettre en évidence les relations qui existent entre le quotient ouvert Q_o et les mécanismes laryngés [RCMAG87, Hen01, HdDC05, RHC09]. De la même manière que pour la fréquence fondamentale, une transition de mécanismes dans le sens MI→MII se caractérise par un saut de Q_o positif et par un saut négatif lors d'une transition dans le sens inverse.

Observations in-vivo

Une grande majorité de la littérature décrivant des observations in-vivo de transition de mécanismes laryngés porte sur la voix chantée [RCMAG87, SSM99, Mil00, Hen01, Lam10]. Les sujets de ces expériences sont des chanteurs avec différents niveaux de technique de chant lyrique. Les protocoles utilisés pour observer ces transitions sont alors essentiellement basés sur la hauteur d'une note ou d'une série de notes. Les consignes données aux sujets sont de réaliser un *glissando* ascendant puis descendant [Rou93, Hen01] ou bien de produire la même note alternativement dans l'un et l'autre des mécanismes [Rou93, Lam10]. Les stratégies d'ajustement, qu'emploient les chanteurs, y compris les amateurs, sont probablement complexes et très finement réalisées. En effet, ces études montrent que l'amplitude mesurée du saut de fréquence lors d'une transition de mécanismes MI et MII diminue lorsque l'expertise des chanteurs augmente.

À l'inverse, Švec et al. [SSM99] demandent à leur sujets de produire une transition volontaire entre les mécanismes MI et MII sans aucune consigne sur la hauteur de la note de départ et sur l'intervalle du saut de fréquence ainsi produit. Certains auteurs observent des transitions entre mécanismes lorsque l'effort pulmonaire au cours de la phonation [SP94, VAL95b] augmente progressivement. L'objectif de ce type de protocole est d'étudier les transitions produites de façon involontaire. Ces études révèlent que les transitions peuvent s'accompagner d'un phénomène de doublement de période [SSM99] ou bien de production vocales diphoniques [VAL95b].

Observations sur larynx excisés

Les mesures réalisées sur des larynx excisés permettent des observations plus aisées des différents phénomènes intervenant dans la phonation que lors des investigations *in vivo*. Elles permettent de plus de contrôler plus précisément et indépendamment les grandeurs physiques

intervenant dans le processus de phonation. Il est tout de même important de noter que ces observations peuvent être biaisées par la modification du comportement physique du larynx, notamment par le fait que celui-ci n'est plus couplé aux résonateurs acoustiques que constituent la trachée et le conduit vocal.

Les premières études caractérisant les différents registres vocaux sur des larynx excisés humains ont été menées par Van den Berg [dB59, dB60, dB68]. En augmentant la tension longitudinale des cordes vocales, il identifie deux régimes d'oscillation distincts, un premier régime nommé « chest » et associé au mécanisme I, observable pour une faible tension et un second régime nommé « falsetto », associé au MII, observé pour une tension plus importante. Il mentionne également que lors d'une augmentation de la tension, l'apparition du MII peut être retardée par un rapprochement des cordes vocales (adduction) ou bien par une augmentation du débit glottique [dB60].

En utilisant des larynx excisés de canins, Berry et al. [BHTS96] observent plusieurs types de régimes d'oscillation en modifiant la pression sous-glottique, la tension longitudinale et le niveau d'adduction des cordes vocales. Parmi les différents régimes observés, les auteurs identifient deux régimes qu'ils associent au mécanismes I et II. Leurs résultats suggèrent notamment que la transition dans le sens MII→MI peut être obtenue en augmentant la pression sous-glottique.

En utilisant des larynx humains excisés, Švec et al. [SSM99] reproduisent des transitions entre deux régimes associés aux mécanismes I et II, en jouant sur la tension longitudinale des cordes vocales. Les auteurs observent que la fréquence fondamentale pour laquelle se produit la transition MI→MII est plus élevée que celle mesurée lors de la transition MII→MI. Cet effet d'hystérésis témoigne de la coexistence de deux régimes pour les mêmes paramètres de contrôle.

Alipour et al. [AFS09] réalisent des mesures de pression sous-glottique, de débit glottique et de signaux EGG sur des larynx canins excisés. Les auteurs observent des transitions de mécanismes, associées à des sauts de fréquence fondamentale en modifiant la tension des cordes vocales, leur niveau d'adduction ou le débit glottique. Ces mesures mettent en évidence la diminution de la résistance glottique au passage de l'air lors de la transition dans le sens MI→MII.

Ces études montrent que les transitions de mécanisme laryngé peuvent être induites par diverses stratégies motrices, impliquant des variations de la tension des cordes vocales, de leur niveau d'adduction ou bien de l'effort pulmonaire.

3.2.2 Hypothèses sur l'origine des transitions de mécanisme

La première hypothèse proposée pour expliquer l'origine des transitions brutales entre mécanismes laryngés est liée à l'activité des muscles laryngés [dB60, HVO70, Tit94]. L'argument principal de cette hypothèse est l'observation par électromyographie de la chute de l'activité du muscle thyro-aryténoïdien (muscle vocal) lors de la transition MI→MII [HVO70]. La contraction du muscle thyro-aryténoïdien atteint alors une limite physiologique. En mécanisme II, la tension des cordes vocales est donc assurée uniquement par l'activité du

muscle crico-thyroïdien qui exerce une tension longitudinale importante sur les cordes vocales, ce qui explique la diminution de leur épaisseur. Selon Alipour et al. [AFS09], ces transitions de mécanismes peuvent se produire avec un larynx passif, sans aucune modification de la tension ou du niveau d'adduction des cordes vocales.

Plusieurs études suggèrent que les interactions acoustiques avec le résonateur situé en amont du larynx (trachée) jouent un rôle majeur dans l'apparition de transitions involontaires entre les mécanismes. Pour Van den Berg [dB68], la transition entre les mécanismes est provoquée par un ajustement musculaire qui se produit pour compenser une interaction acoustique défavorable à l'auto-oscillation. Selon Titze [Tit88a], une interaction favorable entre la première résonance acoustique de la trachée et la vibration laryngée, facilite l'apparition du mécanisme I tandis qu'une interaction défavorable entraîne l'apparition du mécanisme II. Des mesures réalisées sur des chanteurs confirmés [Lam10] et inexpérimentés [TZKH10] montrent que, lors de la réalisation d'un *glissando* ascendant, la fréquence fondamentale pour laquelle se produit la transition dans le sens MI→MII, est plus faible pour les voyelles fermées. Ces résultats impliquent que la fréquence pour laquelle la transition se produit évolue avec la fréquence de la résonance acoustique du conduit vocal. Selon Lamesch [Lam10], cette tendance n'est pas statistiquement significative. D'après Švec [SSM99], le rôle du résonateur acoustique aval sur les transitions entre mécanismes laryngés est secondaire par rapport à celui des propriétés des cordes vocales.

Selon Vilkman et al. [VAL95b], les productions associées au mécanisme I apparaissent lorsque la collision entre les cordes vocales est suffisamment importante dans les dimensions verticales et longitudinales. Le mécanisme II correspond alors au cas où cette condition n'est pas satisfaite.

L'hypothèse la plus récente est la suivante. Les transitions brutales observées lors d'ajustements phonatoires progressifs seraient les manifestations d'une bifurcation du système dynamique non linéaire, qui résulte de l'interaction entre l'écoulement, l'élasticité des cordes vocales et la propagation acoustique. Depuis le début des années 90, les outils de la théorie des systèmes dynamiques non linéaires ont été appliqués à la phonation pour exploiter les modèles théoriques [Luc93, HK95, Luc96, MH97] et pour interpréter les mesures réalisées sur des larynx excisés [BHTS96, SSM99].

3.2.3 Simulations numériques de bifurcations et de transitions de mécanisme laryngé

À l'aide du modèle à deux masses de Ishikawa et Flanagan [IF72], Herzel et Knudsen [HK95] simulent des bifurcations aux frontières de la région de paramètres correspondant à des conditions de phonation typiques de la voix parlée. Ils observent plusieurs types de bifurcations associées à des transitions vers des régimes d'oscillations quasi-périodiques et chaotiques. L'apparition de ce type de régime est interprétée comme une manifestation de la co-existence de plusieurs cycle limites pour un même jeu de paramètres de contrôle.

Avec le même modèle, Lucero [Luc93] montre que les deux masses peuvent osciller autour de deux couples de positions d'équilibres différents, correspondants à deux régions différentes

de paramètres de contrôle¹. Dans une autre étude [Luc96], le même auteur exploite cette propriété pour simuler des oscillations comparables à celles des mécanismes laryngés I et II. Le régime d'oscillation associé au MII est obtenu en augmentant la valeur de la raideur k_2 de la deuxième masse. La position d'équilibre ainsi obtenue correspond à une ouverture glottique plus large et plus convergente. La différence entre les deux régimes d'oscillation se caractérise également par le déphasage entre le mouvement des deux masses, qui est plus faible pour le régime associé au mécanisme MII. L'auteur mentionne également que le phénomène de co-existence de deux cycles limites n'apparaît pas dans les simulations réalisées avec le modèle de Ishikaza et Flanagan [IF72] en tenant compte de la viscosité de l'écoulement ou bien de la formation d'un jet au niveau de la constriction glottique.

En utilisant le modèle à deux masses de Lous [LHVH98], Sciamarella et d'Alessandro [SD04] s'intéressent à l'évolution des paramètres acoustiques des simulations (fréquence fondamentale, quotient ouvert ...) en fonction des paramètres du modèle. Dans leur analyse, les auteurs distinguent deux régimes d'oscillation différents : avec et sans fermeture complète de la glotte. Ces régimes sont associés respectivement au mécanismes laryngés I et II. Cette étude montre que ce modèle permet de reproduire les sauts de fréquence provoqués par la transition entre les deux mécanismes laryngés principaux. De plus, leurs résultats suggèrent qu'il est possible d'observer cette transition en faisant varier uniquement un paramètre de contrôle parmi les suivants : $(P_{sub}, l_g, L_g, k_i, k_c)$.

Contrairement aux études décrites ci-dessus, qui ne tiennent pas compte du couplage acoustique, Mergel et Herzel [MH97] s'intéressent au rôle du conduit vocal dans l'apparition d'une instabilité vocale et en particulier au phénomène de bi-phonation (deux composantes fréquentielles indépendantes) dans le cas d'une fréquence fondamentale élevée. Ils utilisent pour cela un modèle à deux masses asymétrique, couplé à un résonateur aval de section uniforme. Leur conclusion indique que les instabilités vocales apparaissent avec et sans prise en compte du couplage acoustique. Cependant, les simulations réalisées en considérant le conduit vocal montrent que ces instabilités vocales sont étendues à une région de paramètres de contrôle plus importante et notamment pour une pression sous-glottique plus faible.

Malgré les nombreuses études sur la voix chantée cherchant à caractériser la transition entre les mécanismes I et II et le saut de fréquence qui l'accompagne [RCMAG87, SSM99, Lam10], la reproduction de ce type de transition dynamique à l'aide d'un modèle simplifié de cordes vocales n'est pas une approche très répandue. C'est dans cette optique que Tokuda et al. [THvH07] utilisent un modèle comprenant trois masses alignées, sans résonateur acoustique. Les bifurcations vers le mécanisme II sont atteintes en augmentant la raideur de la masse centrale. Les simulations réalisées mettent en évidence que ces bifurcations sont soumises à un phénomène d'hystérésis qui témoigne de la co-existence de deux cycles limites possibles pour certains jeux de paramètres de commande. Les auteurs interprètent les mécanismes laryngés comme différents modes propres de la structure mécanique des cordes vocales, ayant chacun leur fréquence propre et leur motif de vibration. Les changements de mécanismes sont expliqués par une modification de la prédominance des modes propres.

Dans une autre étude [TZKH10], les mêmes auteurs s'intéressent à l'effet du couplage acoustique sur les changements involontaires de mécanisme laryngé. À l'aide d'un modèle

1. En se référant à la théorie des système dynamique [GH83], Lucero précise que les deux régions de paramètres sont séparées par une bifurcation trans-critique

comprenant quatre masses, dont trois sont alignées et représentent la couche superficielle des cordes vocales, les auteurs simulent des *glissandi* en augmentant la valeur du paramètre de tension des quatre masses (augmentation de la raideur et diminution de la masse). Le mécanisme I est associé à la collision des trois masses représentant la couche superficielle tandis que la mécanisme II est associé au cas où le nombre de masses qui s'entrechoquent est inférieur à trois. Leur conclusion est que les résonateurs amont et aval ont une influence considérable sur la fréquence fondamentale pour laquelle le changement de régime se produit. En effet, leurs simulations prédisent que la valeur de la fréquence de transition augmente lorsque les fréquences de résonance acoustique augmente.

3.3 Observation *in vivo* de saut de fréquence fondamentale

3.3.1 Matériel et méthode

L'objectif de cette section est d'observer des sauts de fréquence fondamentale produits de façon involontaire. Pour ce faire, des mesures acoustiques et électroglottographiques sont réalisées sur un locuteur souffrant de pertes auditives importantes.

3.3.1.1 Dispositif expérimental

Le dispositif expérimental utilisé pour cette étude est la station EVA 2 [TG95, TGG99, GYR⁺06], développée au laboratoire Parole et Langage d'Aix-en-Provence. La station EVA permet de réaliser simultanément des mesures acoustiques, aérodynamiques et électroglottographiques durant la production de parole. Les signaux des différents capteurs sont conditionnés puis numérisés de façon synchrone par une unique carte d'acquisition. La station EVA est présentée plus en détails dans l'annexe B qui traite de la calibration des capteurs aérodynamiques et des problèmes de synchronisation liés à l'utilisation de la station. Dans cette étude, les capteurs utilisés sont le microphone intégré à la station EVA (AKG, modèle C419) et un électroglottographe (Laryngograph) connecté à la station. Les signaux acoustique et électroglottographique sont échantillonnés à une fréquence de 25 *kHz*.

3.3.1.2 Sujet et protocole

Sujet

Les productions vocales enregistrées ont été produites par un locuteur malentendant, appelé sujet Y. Le sujet Y est un homme de 60 ans souffrant d'une surdité de transmission², survenue dans l'enfance après l'acquisition du langage (surdité post-linguistique). Le sujet Y

2. Surdité liée à un dysfonctionnement de l'oreille externe ou moyenne.

est équipé de deux types de prothèse auditive, appareil auditif traditionnel qui amplifie la conduction aérienne dans l'oreille externe (composé d'un microphone, d'un amplificateur et d'un haut-parleur) et un implant cochléaire activé depuis 3 semaines au moment des mesures. Afin de restreindre le contrôle moteur liée la perception auditive du sujet, ses deux prothèses auditives ont été désactivées quelques minutes avant la réalisation des mesures.

Protocole

La consigne donnée au participant est de produire les trois voyelles cardinales : /a/, /u/ et /i/ et de maintenir la phonation pendant 5 secondes minimum, pour 3 niveaux d'intensités différents :

- avec une intensité constante importante, notée « fort »
- avec une intensité constante plus faible, notée « faible »
- en augmentant progressivement l'intensité au cours de la production, notée « *crescendo* ».

Aucune consigne concernant la hauteur tonale de la production n'est spécifiée.

3.3.1.3 Méthode d'analyse des signaux

Détermination du quotient ouvert

La détermination du quotient ouvert Q_o nécessite d'estimer les instants de fermeture t_f et d'ouverture t_o glottique. Cette tâche, réalisée automatiquement à partir des signaux électroglottographiques, est décrite par les étapes suivantes.

- ① Pour chaque période d'indice n , l'instant $t_z[n]$ correspondant au passage par zéro du signal *EGG* est déterminé (lorsque le signal devient négatif). Cette opération est réalisée sur le signal *EGG* filtré passe-haut, à l'aide d'un filtre de Butterworth d'ordre 4 et de fréquence de coupure 80 *Hz*, afin de supprimer les perturbations dues aux fluctuations basses fréquences du signal.
- ② L'instant de fermeture $t_f[n]$ est déterminé à partir de l'instant correspondant au maximum du signal *DEGG* dans l'intervalle compris entre $t_z[n - 1]$ et $t_z[n]$.
- ③ L'instant $t_o[n]$ correspondant à l'ouverture des cordes vocales est estimé partir du minimum du signal *DEGG* dans l'intervalle compris entre $t_f[n]$ et $t_f[n + 1]$.

La figure 3.1 illustre la détection de $t_f[n]$ et $t_o[n]$ dans le cas d'une production vocale en mécanisme I.

La détermination des instants t_f et t_o permet ensuite de calculer, pour chaque cycle de vibration glottique d'indice n , la valeur du quotient ouvert Q_o , défini comme le rapport de la durée de la phase ouverte sur la durée de la période de la phonation tel que

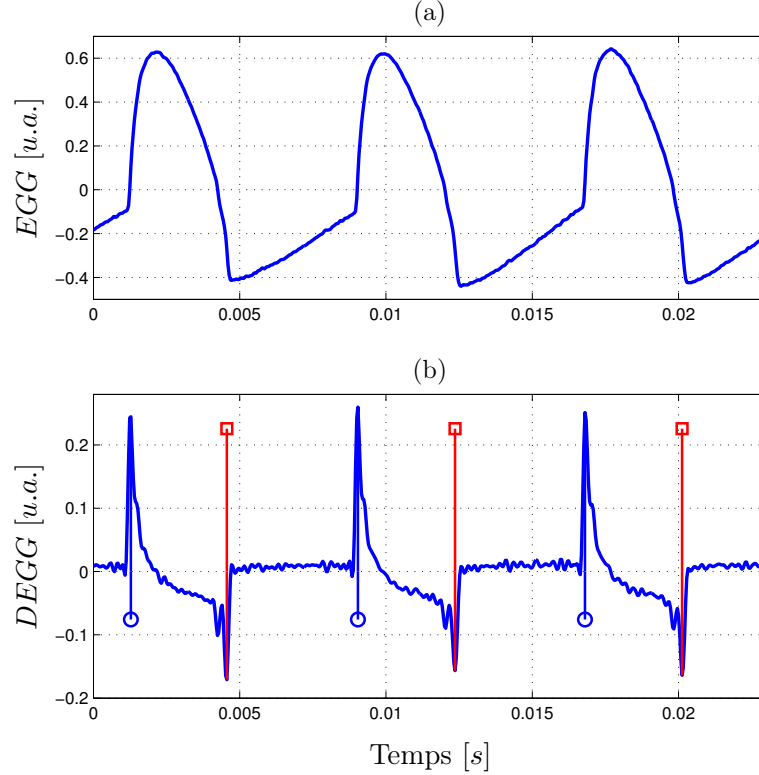


FIGURE 3.1 – (a) : Signal EGG pour la voyelle /a/ produite en mécanisme I par un sujet masculin normo-entendant. (b) : DEGG, dérivée temporelle du signal EGG. L’amplitude des signaux est donnée en unité arbitraire (*u.a.*). Les marqueurs représentent les instants t_f (○) et t_o (◻) détectés.

$$Qo[n] = \frac{t_f[n+1] - t_o[n]}{t_f[n+1] - t_f[n]}. \quad (3.1)$$

Dans le cas présenté en figure 3.1, le quotient ouvert est de l’ordre de 60 %.

Détermination de la fréquence fondamentale

L’évolution de la fréquence fondamentale du signal de pression acoustique est estimée en utilisant méthode YIN [dCK02]. Cette méthode est basée sur l’autocorrélation de la forme d’onde temporelle du signal. La périodicité du signal est estimée à partir d’un seuil sur l’amplitude de l’autocorrélation. La résolution temporelle de l’estimation est de 1,28 ms.

3.3.2 Résultats

L’analyse de l’ensemble des productions vocales fait apparaître deux groupes. Les voyelles fermées (/i/ et /u/) pour lesquelles $300 \leq F_0 \leq 600$ Hz et $50 \leq Q_o \leq 80$ % et les voyelles ouvertes (/a/) pour lesquelles $100 \leq F_0 \leq 400$ Hz et $30 \leq Q_o \leq 50$ %.

La figure 3.2 montre un exemple de production de la voyelle /u/. Les signaux expérimentaux révèlent l'apparition d'un phénomène d'instabilité vocale, pour $0,1 < t < 0,3$ s, qui se manifeste par un phénomène de battement visible sur l'enveloppe temporelle du signal, une modification du contenu spectral et par une variance du quotient ouvert estimé, localement plus importante. L'apparition de l'instabilité est précédée par une diminution de F_0 . Le deuxième événement important est une diminution brusque de F_0 à $t = 0,7$ s, qui passe de 500 à 390 Hz en 30 ms. Cette chute est associée à une discontinuité importante dans l'évolution du quotient ouvert Q_o dont la valeur varie quasiment instantanément de 70 à 39 %. Le spectrogramme montre que cet événement intervenant à la fin de la production se caractérise également par un enrichissement du contenu spectral du signal de pression p_a .

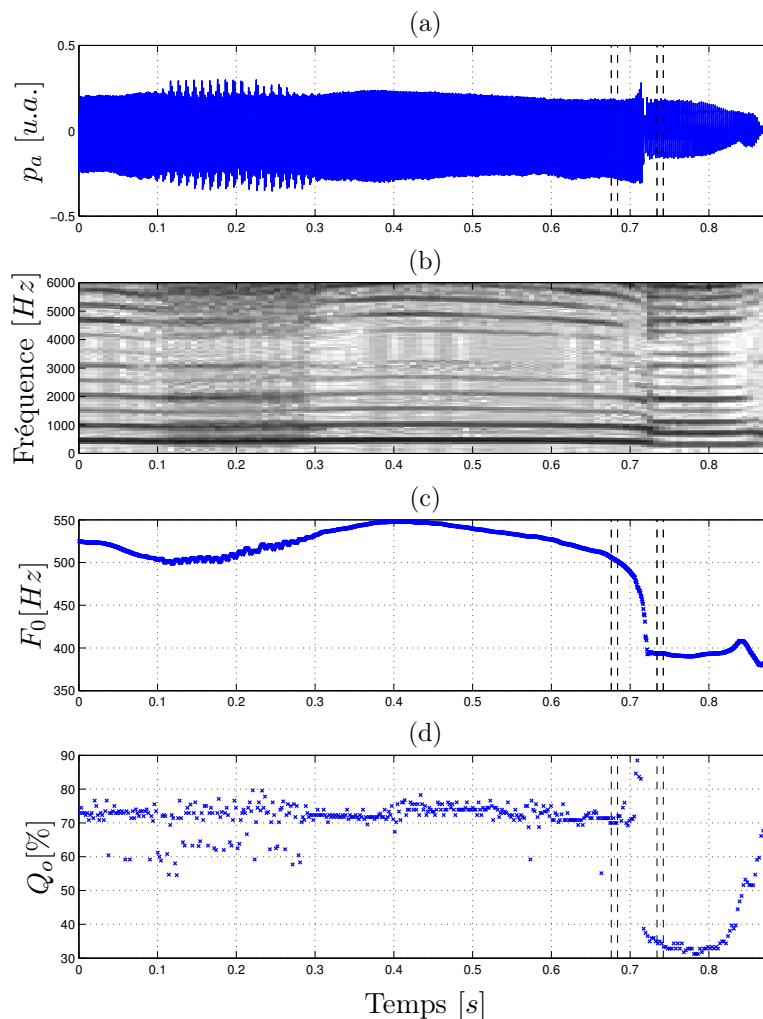


FIGURE 3.2 – Voyelle /u/ produite avec une forte intensité par le sujet Y. (a) : Signal de pression acoustique p_a mesurée. (b) : Spectrogramme de p_a calculé avec des fenêtres temporelles de 10 ms. (c) : Fréquence fondamentale F_0 du signal p_a . (d) : Quotient ouvert Q_o extrait du signal EGG. Les traits pointillés verticaux délimitent l'intervalle de temps choisi pour la représentation des signaux EGG en figure 3.3. [Ecouter le signal \$p_a\$ \(son 1\)](#)

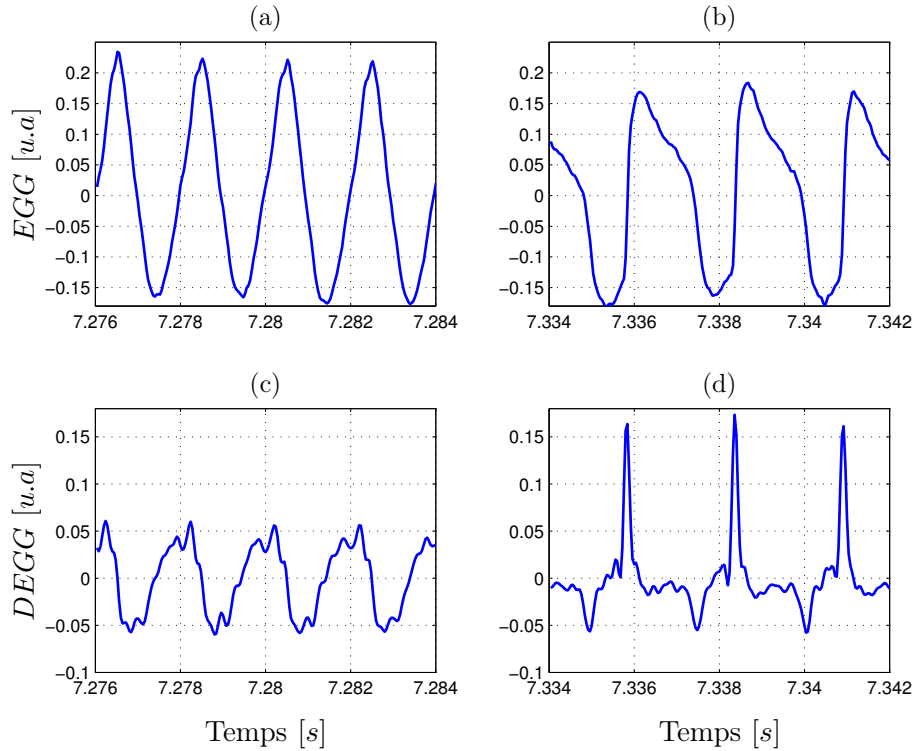


FIGURE 3.3 – Signaux *EGG* (a,b) et *DEGG* (c,d) mesurés pour la voyelle /u/ produite avec une forte intensité par le sujet Y, avant (a,c) et après (b,d) le saut de fréquence fondamentale observé en figure 3.2.c.

La figure 3.3 montre que cet événement induit une modification de la forme d'onde du signal *EGG*. La forme d'onde du signal *EGG* devient plus asymétrique après le saut de fréquence fondamentale. L'analyse des signaux *DEGG* révèle également l'apparition de pics positifs dans la forme d'onde qui suggèrent qu'après le saut de fréquence, la surface de contact entre les deux cordes vocales augmente plus rapidement lors de la collision.

Bien qu'avant le saut de fréquence fondamentale le son produit soit bien perçu comme la voyelle /u/, le son semble plus proche des voyelles /ɔ/ et /ɑ/ après cet événement. Cette perception est confirmée par une estimation de la valeur des fréquences de résonance du conduit vocal par prédiction linéaire, réalisée avec 30 pôles. Cette estimation donne pour $0,72 \leq t \leq 0,81$ s et pour les deux premières résonances, $F_1 \simeq 680$ Hz et $F_2 \simeq 1050$ Hz. Avant le saut de fréquence fondamentale, l'estimation de F_1 n'est pas exploitable en raison de la valeur élevée de F_0 .

D'autres cas de sauts de fréquence fondamentale sont observés lors de la production d'une voyelle fermée prononcée avec une intensité forte ou *crescendo*. Certains d'entre-eux sont présentés en annexe C pour le sujet Y et pour un autre sujet sourd.

3.3.3 Discussion et bilan

Discussion

Dans un premier temps, les mesures décrites dans cette section montrent que les fréquences fondamentales des voyelles produites par le sujet Y sont en moyenne plus élevées que les valeurs typiques rencontrées chez les hommes normo-entendants, situées entre 100 et 150 Hz [Meu07]. Ces mesures confirment les résultats des études expérimentales menées sur la production de la parole chez les personnes sourdes [AKH64, Boo66, Mar68]. L'écart constaté entre les valeurs de F_0 pour les voyelles fermées et ouvertes suggère que les productions du sujet Y sont marquées par une accentuation de l'effet de la fréquence fondamentale intrinsèque des voyelles (cf. section 1.3.2). De plus, les valeurs de Q_0 supposent que les voyelles ouvertes sont systématiquement produites en mécanisme I et que les voyelles fermées sont produits en mécanisme II.

Les chutes brusques des valeurs de F_0 et de Q_0 pour les voyelles fermées, produites avec une forte intensité témoignent d'une transition de mécanismes laryngés dans le sens MII→MI. Cette interprétation est corroborée par l'évolution de la forme temporelle du signal *EGG* dont l'asymétrie augmente lors du passage en MI [CL91, HRC+03].

Les sauts de fréquence fondamentale sont observés lors de la production de voyelles fermées, pour lesquelles la fréquence F_1 de la première résonance du conduit vocal est relativement basse (environ 300 Hz). Dans ce cas, le faible écart entre F_1 et F_0 rend difficile l'estimation de la valeur de F_1 à partir du signal de pression acoustique. Après le saut de fréquence, les fréquences F_1 et F_2 estimées correspondent à celles d'une voyelle plus ouverte, ce qui suppose qu'un mouvement d'ouverture de la constriction formée entre le palais et la langue est réalisé par le locuteur. Une perspective intéressante serait de reproduire ces expériences en mesurant l'impédance ou la fonction de transfert du conduit vocal [ESW97, JSW04] pendant la production afin de disposer d'informations fiables sur les résonances acoustiques du conduit vocal qu'elle que soit la valeur de la fréquence fondamentale.

Bilan

Les mesures réalisées ont permis d'observer des sauts de fréquence fondamentale produits de façon involontaire par un locuteur ne disposant d'aucun retour auditif lors de la production. Les signaux *EGG* montrent que ces sauts de fréquence sont les manifestations de transitions involontaires entre mécanismes laryngés. L'analyse des signaux acoustiques suggère que le saut de fréquence observé est associé à un mouvement d'ouverture de la constriction qui résulte d'un abaissement de la langue. Les travaux menés dans cette deuxième partie du manuscrit ont pour objectifs de déterminer les stratégies motrices responsables de l'apparition de ces transitions ainsi que leurs influences sur l'écart entre les valeurs de fréquence fondamentale mesurées avant et après le saut.

Validation expérimentale des modèles théoriques par simulation numérique

Sommaire

4.1	Introduction	69
4.2	Dispositif expérimental	70
4.2.1	Maquette de l'appareil phonatoire	70
4.2.2	Capteurs de mesure	70
4.2.3	Maquette de cordes vocales	71
4.3	Résultats expérimentaux	73
4.3.1	Variation de la pression d'eau	73
4.3.2	Variation de la pression d'alimentation	75
4.3.3	Variation de la longueur du résonateur aval	77
4.4	Simulation numérique	77
4.4.1	Choix des paramètres du modèle à deux masses	78
4.4.2	Résultats	80
4.5	Discussion et conclusion	87
4.5.1	Discussion	87
4.5.2	Conclusion	90

4.1 Introduction

L'objectif de ce chapitre est d'évaluer la capacité des modèles théoriques présentés au chapitre 2 à reproduire les transitions de mécanismes laryngés.

Dans cette optique, des mesures sont réalisées sur une maquette auto-oscillante de l'appareil phonatoire. Les mesures et les simulations numériques font apparaître des transitions entre deux régimes d'oscillation observées en faisant varier trois paramètres de contrôle différents. Ces transitions sont accompagnées de sauts de la fréquence fondamentale et dans certains cas d'un phénomène d'hystérésis. L'évaluation des modèles est basée sur les valeurs des paramètres de contrôle et de la fréquence fondamentale lors de la transition.

4.2 Dispositif expérimental

La maquette présentée dans cette section a vocation à reproduire le fonctionnement physique de la production d'un son voisé sans chercher à imiter le détail de l'anatomie humaine. Par rapport aux observations *in vivo*, l'utilisation de maquette améliore le contrôle et l'observabilité des expériences, ce qui facilite l'étude des différents phénomènes physiques.

4.2.1 Maquette de l'appareil phonatoire

Les modèles théoriques présentés dans le chapitre 2 sont évalués sur une maquette de l'appareil phonatoire constituée des éléments suivants.

- Un réservoir de pression (4) de $0,6 \text{ m}^3$ de volume, alimenté par un compresseur (1) (Atlas Copco, GA5 FF-300-8) et régulé par un détendeur (3) (Norgren, 11-818-987), pour représenter les poumons. Du fait de ses dimensions importantes, le réservoir agit comme une alimentation en pression quasi-constante. Le réservoir est également rempli de mousse acoustique de façon à limiter l'effet des modes acoustiques de cavité.
- Un résonateur amont (6) pour représenter la trachée, constitué d'un tube uniforme en duralium, de diamètre interne $2,5 \text{ cm}$ et de longueur $20,5 \text{ cm}$ dont environ 10 cm à l'intérieur du réservoir de pression.
- Une maquette de cordes vocales auto-oscillante (7), composée de deux tubes en latex remplis d'eau.
- Un résonateur aval (8) pour représenter le conduit vocal, constitué d'un tube uniforme en duralium, de diamètre interne $2,5 \text{ cm}$ et un tube en plexiglas de longueur variable.
- Un réservoir d'eau (5), placé au-dessus du niveau de la maquette de cordes vocales, qui communique avec les volumes d'eau à l'intérieur des répliques de cordes vocales.

La figure 4.1 illustre le dispositif.

4.2.2 Capteurs de mesure

Un capteur de pression différentielle (Kulite, XCS-093-035-Bar-G) est utilisé pour mesurer la pression statique, P_0 en amont de la maquette de cordes vocales. Ce capteur est calibré au moyen d'un manomètre à eau avec une précision de $\pm 5 \text{ Pa}$. Un microphone (Bruel & Kjaer, 2265), placé à 20 cm au dessus de la réplique de cordes vocales, mesure la pression acoustique p_{ac} . Les signaux issus de ces deux capteurs sont enregistrés au moyen d'une carte d'acquisition (National Instruments, PCI-MIO-16XE-10) à une fréquence d'échantillonnage de 10 kHz via le logiciel Labview.

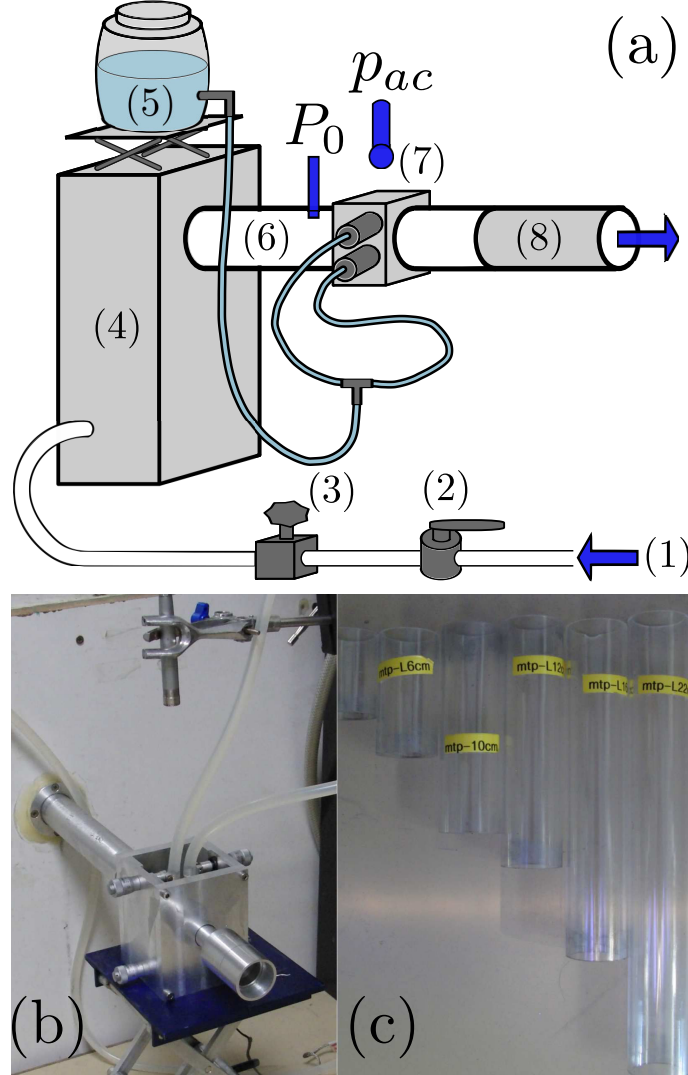


FIGURE 4.1 – (a) : Schéma de la maquette de l'appareil phonatoire, (1) : compresseur, (2) : vanne à boisseau, (3) : détendeur, (4) : réservoir de pression, (5) : réservoir d'eau, (6) : résonateur amont, (7) : réplique de cordes vocales, (8) : résonateur aval incluant un tubes en plexiglas. (b) : Photographie de la maquette. (c) : Photographie des différents tubes en plexiglas utilisés, de longueur 4, 6, 10, 12, 16 et 22 *cm*.

4.2.3 Maquette de cordes vocales

La maquette de cordes vocales utilisée ici est inspirée de celles développées pour représenter les lèvres des joueurs de cuivres [GPP98, PG99, CGC00] et est directement issue des maquettes de cordes vocales développées à Grenoble [Vil02, Rut07, Bai09, Her14]. Toutes ces répliques sont basées sur l'utilisation de deux tubes en latex remplis d'eau qui entrent en auto-oscillation lorsqu'ils sont traversés par un écoulement d'air suffisamment important. Les différents éléments de cette maquette sont détaillés en figure 4.2.

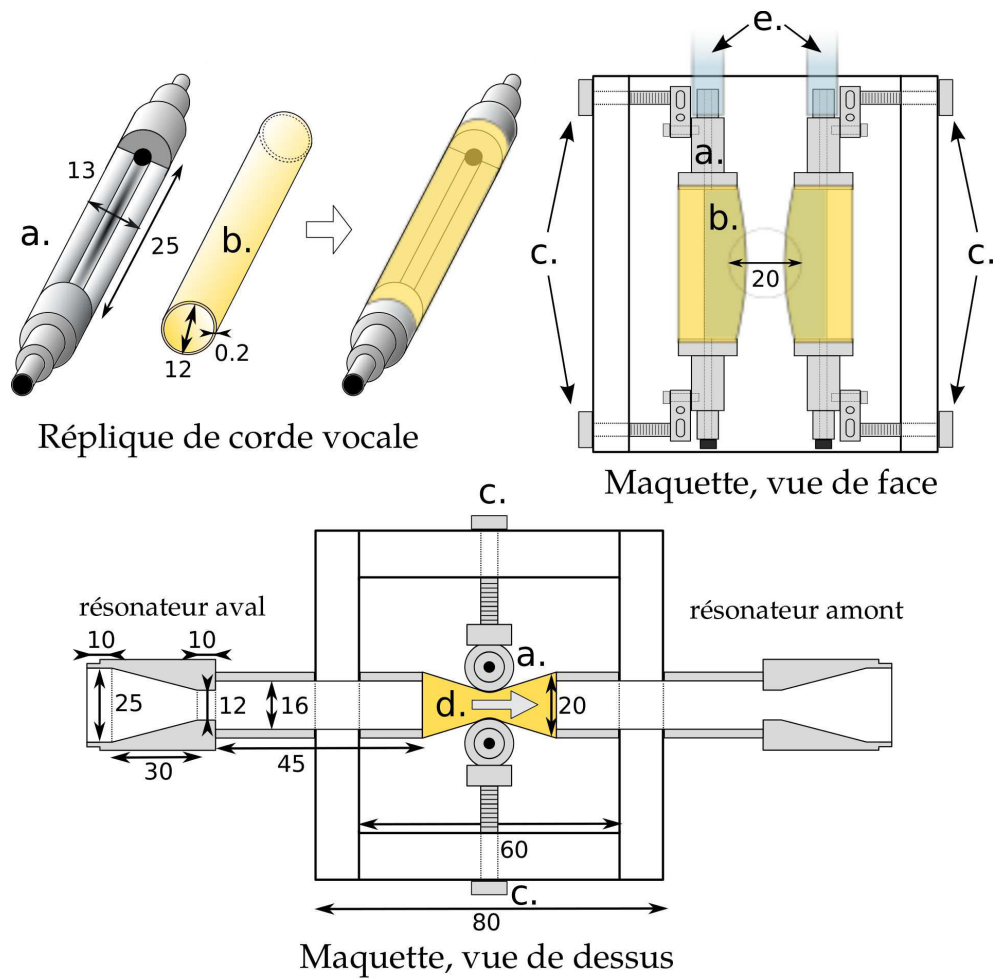


FIGURE 4.2 – Schéma de la maquette de cordes vocales. Chaque réplique de corde vocale est constituée d'un support semi-cylindrique en métal (a.), entouré par un tube en latex (b.) rempli d'eau. Des vis micrométriques (c.) permettent de régler l'écartement entre les deux répliques de corde vocale. L'écoulement d'air est guidé à travers la constriction formée par les cordes vocales par l'intermédiaire d'un boyau central en latex (d.) d'épaisseur $0,3\text{ mm}$ et de diamètre interne 16 mm . Deux tubes de diamètre 7 mm relient le volume d'eau contenue dans les cordes vocales au réservoir d'eau (e.). Les dimensions sont indiquées en millimètres.

La spécificité de cette maquette par rapport aux précédentes [Vil02, Rut07, Bai09, Her14] réside dans la présence des deux couches de latex formées par les tubes et le boyau central, afin d'imiter les différentes couches de muqueuse qui constituent les cordes vocales humaines, et la possibilité de régler finement l'écartement entre les répliques de corde vocale.

La configuration de la maquette de cordes vocales est décrite par trois paramètres réglables :

- l'écartement entre les répliques de corde vocale, réglable à l'aide des vis micrométriques,

- la pression statique de l'eau dans les cordes vocales
- la tension axiale du boyau central en latex qui guide l'écoulement.

4.3 Résultats expérimentaux

L'objectif est d'observer des sauts de fréquence provoqués par une bifurcation du système auto-oscillant, c'est-à-dire une variation brusque de la fréquence fondamentale F_0 lors d'une variation progressive de l'un des paramètres de contrôle de la maquette. La quantité observée est donc F_0 et les paramètres de contrôles utilisés dans cette étude sont :

- la pression d'eau P_e à l'intérieur des tubes en latex, qui dépend de la hauteur h_e du réservoir d'eau par rapport à la maquette de cordes vocales, telle que $P_e = \rho g h_e$, h_e est modifiée manuellement par l'expérimentateur,
- la pression aérodynamique P_0 modifiée à l'aide de la molette de contrôle du détendeur,
- la fréquence F_1 de la première résonance du résonateur situé en aval de la maquette de cordes vocales dont la longueur totale est modifiée en utilisant différents tubes en plexiglas de longueur L_p .

L'écartement entre les répliques de corde vocale ainsi que la tension du boyau central en latex restent inchangés durant toutes les mesures présentées dans cette section. Les mesures sont réalisées à partir d'une configuration de référence en faisant varier chaque paramètre de contrôle séparément. Les paramètres de la configuration de référence sont donnés dans le tableau 4.1. La valeur de F_1 correspond à la fréquence du premier pic de l'impédance d'entrée du résonateur (cf. Annexe D.1), estimée à partir de la géométrie du résonateur aval, à l'aide de l'équation 2.55. Pour la première résonance du résonateur amont, le calcul donne 230 Hz .

Paramètres	P_e	P_0	F_1	F_0
Valeurs	$8000 \pm 5 \text{ Pa}$	$1300 \pm 25 \text{ Pa}$	$830 \pm 10 \text{ Hz}$	$220 \pm 2 \text{ Hz}$

TABLE 4.1 – Valeurs des paramètres de la maquette de l'appareil phonatoire pour la configuration de référence. Cette configuration correspond au cas où la longueur du résonateur aval est la plus faible, c'est-à-dire sans tube en plexiglas ($L_p = 0$), comme sur la figure 4.1.b

4.3.1 Variation de la pression d'eau

À partir de la configuration de référence, la pression d'eau est diminuée progressivement entre les instants $t_1 = 0,5 \text{ s}$ et $t_2 = 8,4 \text{ s}$ jusqu'à atteindre $P_e = 4500 \text{ Pa}$. La pression P_e est ensuite augmentée jusqu'à sa valeur initiale, atteinte à l'instant $t_3 = 22,0 \text{ s}$. Entre ces instants particuliers, l'évolution de P_e est supposée linéaire. La figure 4.3 présente les signaux expérimentaux.

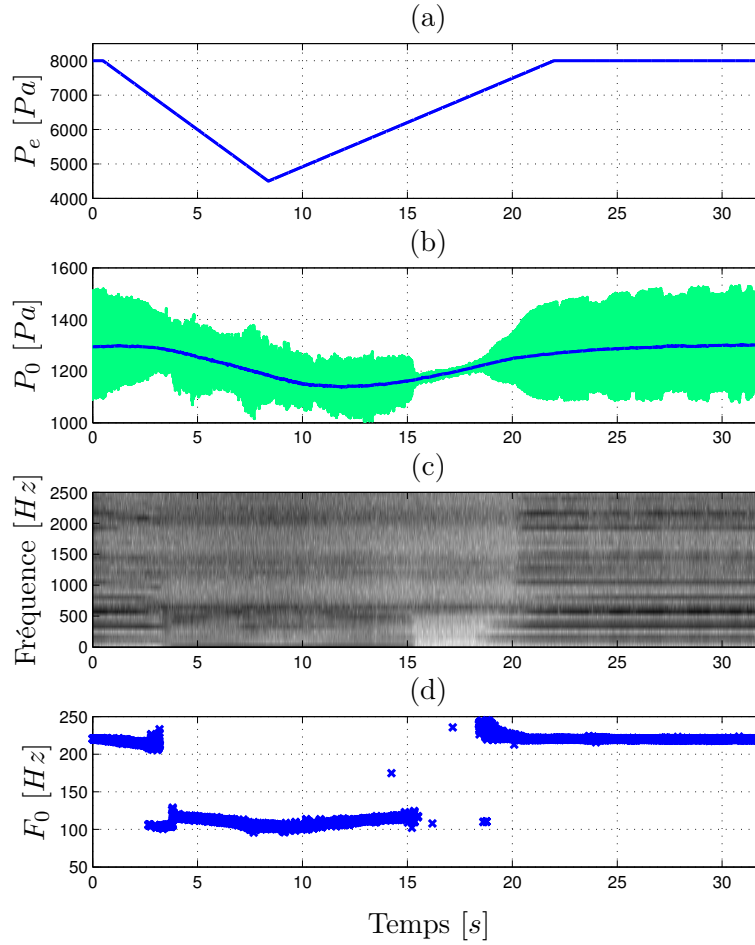


FIGURE 4.3 – (a) : Pression d'eau P_e contenue dans les répliques de cordes vocales. (b) : Pression P_0 mesurée en amont de la maquette de cordes vocales, présentée avec et sans filtrage passe-bas d'ordre 5 et de fréquence de coupure 30 Hz . (c) : Spectrogramme du signal de pression acoustique p_{ac} , calculé avec des fenêtres de Hamming de durée 30 ms . (d) : Fréquence fondamentale F_0 du signal p_{ac} , estimée avec la méthode YIN [dCK02].

[Ecouter le signal \$p_{ac}\$ \(son 2\)](#)

Pendant les deux premières secondes, la diminution de P_e entraîne la diminution progressive de la fréquence fondamentale F_0 de l'auto-oscillation. Ensuite à $t = 3$ s, F_0 chute brusquement de 210 à 120 Hz et continue de diminuer progressivement jusqu'à atteindre un minimum de 100 Hz à $t = t_2$. La valeur de F_0 remonte ensuite progressivement avec P_e jusqu'à ce que l'auto-oscillation s'arrête, à $t = 15$ s. Les auto-oscillations de la maquette reprennent à $t = 18$ s à une fréquence de 240 Hz . Le spectrogramme montre que les sauts de fréquence marquent une transition entre deux phases qui diffèrent par leur densité d'énergie spectrale, en particulier pour des fréquences supérieures à 1 kHz .

Il est important de préciser qu'avec cette maquette de cordes vocales, diminuer P_e fait indirectement diminuer P_0 , avec un retard de l'ordre 3 s. En effet, la pression P_e est responsable du gonflement des tubes en latex, une diminution de P_e entraîne donc une augmentation de l'aire de la constriction glottique [Her14] et donc une diminution de la résistance au passage

de l'air.

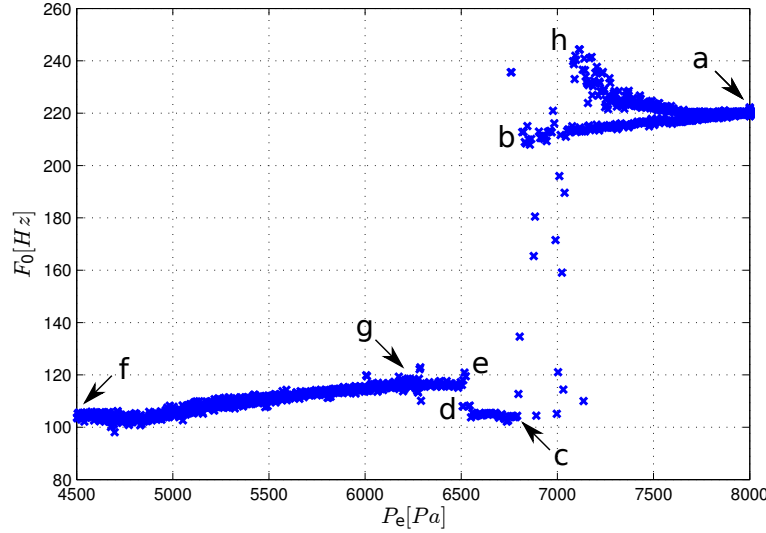


FIGURE 4.4 – Fréquence fondamentale F_0 en fonction de la pression P_e . Les points importants sont repérés par des lettres. L'ordre alphabétique correspond à l'ordre chronologique des événements.

La figure 4.4 décrit l'évolution de F_0 en fonction de P_e lors de la diminution puis de l'augmentation de P_e . Cette représentation permet de caractériser les deux régimes d'oscillation observés : le régime B (RB) qui correspond à la configuration de référence, pour $P_e > 6750 \text{ Pa}$ et $F_0 \geq 210 \text{ Hz}$ et le régime A (RA) pour $P_e < 6500 \text{ Pa}$ et $F_0 \leq 120 \text{ Hz}$. La transition RB→RA est réalisée en passant par la branche **cd** qui correspond à un doublement de période par rapport au point **b**. Lors de la transition dans le sens inverse (RA→RB), l'auto-oscillation est stoppée entre les points **g** et **h**, sur une plage de P_e allant de 6250 à 7000 Pa . Les oscillations reprennent en RB au niveau du point **h** à une fréquence F_0 , supérieure à celle de la branche **ab** pour la même valeur de P_e . Finalement la fréquence F_0 retrouve sa valeur initiale, au point **a**, pour $P_e = 8000 \text{ Pa}$.

Une asymétrie entre les trajectoires associées à la diminution et à l'augmentation de P_e est observée. La disparition de l'auto-oscillation lors de la transition (RA→RB) peut s'expliquer par les faibles valeurs de P_0 lors de cette transition. En ne tenant pas compte des phases correspondantes à un doublement de période ou une aphonie (branches **cd** et **gh**), l'amplitude du saut de fréquence fondamentale est de 94,9 Hz dans le sens RB→RA et de 106,9 Hz dans le sens inverse.

4.3.2 Variation de la pression d'alimentation

En partant de la configuration de référence, la pression P_0 est augmentée au cours de la phonation jusqu'à une valeur de 1650 Pa à l'aide de la molette de contrôle du détendeur,

puis diminuée jusqu'à sa valeur initiale. La figure 4.5 présente les signaux expérimentaux correspondant. Lors de l'augmentation de P_0 , la fréquence F_0 est quasi-constante jusqu'à $t = 22$ s, puis F_0 passe de 220 à 125 Hz. La valeur de F_0 reste stable lorsque P_0 atteint un maximum, à $t = 30$ s, puis augmente pour atteindre la valeur de 200 Hz à $t = 42$ s, lors de la diminution de P_0 . Les transitions associées à ces sauts de fréquence sont clairement visibles sur la forme d'onde du signal P_0 ainsi que sur le spectrogramme de p_{ac} .

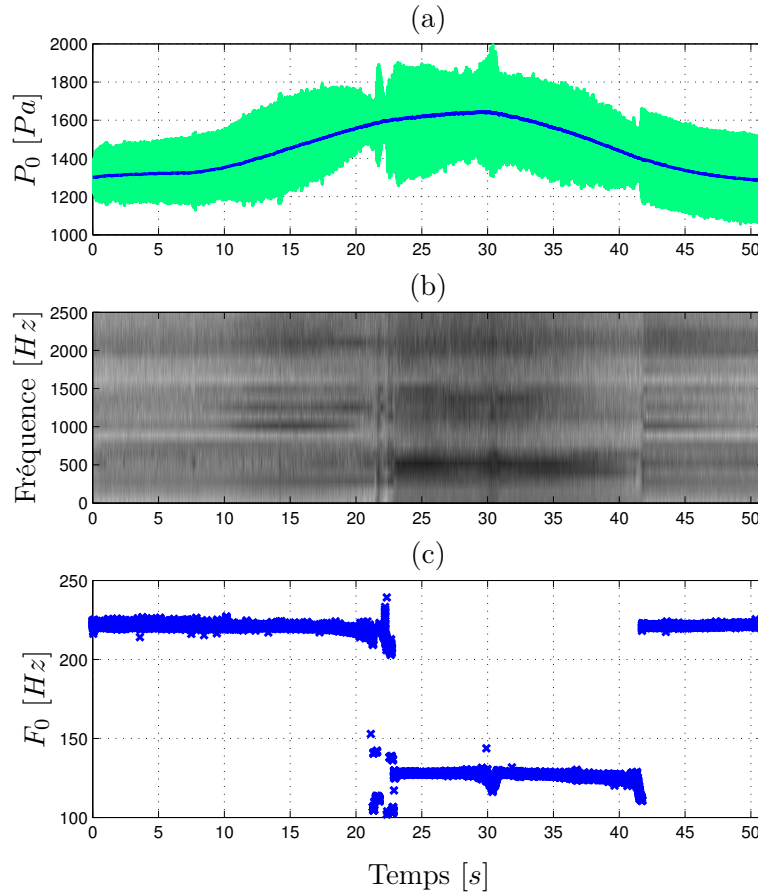


FIGURE 4.5 – (a) : Pression P_0 mesurée en amont de la maquette de cordes vocales, présentée avec et sans filtrage passe-bas d'ordre 5 et de fréquence de coupure 30 Hz. (b) : Spectrogramme du signal de pression acoustique p_{ac} , calculé avec des fenêtres de Hamming de durée 30 ms. (c) : Fréquence fondamentale F_0 du signal p_{ac} . [Ecouter le signal \$p_{ac}\$ \(son 3\)](#)

La figure 4.6 présente la même mesure dans le plan (F_0, P_0) . Il est possible d'identifier à nouveau les deux régimes, RB pour des fréquences F_0 proches de 220 Hz et RA pour $120 < F_0 < 130$ Hz. Durant l'augmentation de P_0 (branche **ad**), la transition RB→RA, entre les points **b** et **c**, se fait de manière chahutée au niveau de l'évolution de F_0 et de l'enveloppe temporelle du signal acoustique. Dans ce cas, aucun doublement de période n'est observé même si l'estimateur indique quelques valeurs de F_0 inférieures à celles décrites par la branche **cd**. La figure 4.6 fait apparaître clairement un cycle d'hystérésis décrit par les points **b**, **c**, **e** et **f**, pour $1400 < P_0 < 1600$ Pa. L'amplitude des sauts de fréquence est de 83,2 Hz lors de la transition RB→RA et de 97,4 Hz lors de la transition dans le sens RA→RB.

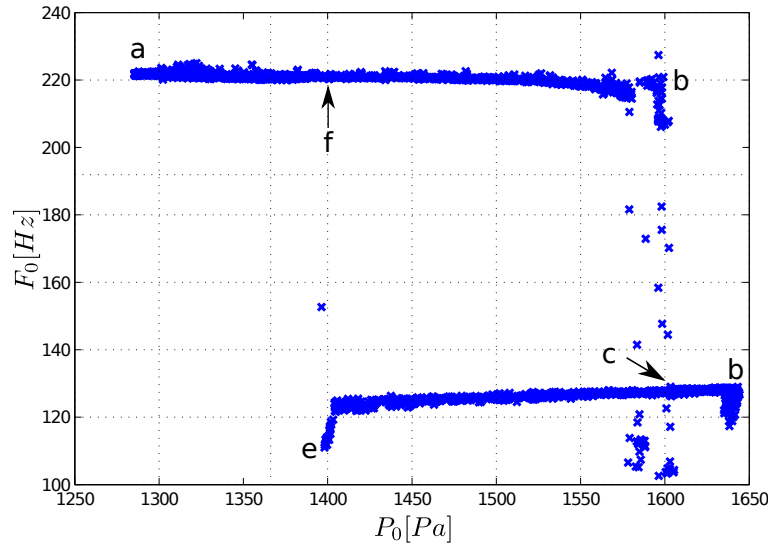


FIGURE 4.6 – Fréquence fondamentale F_0 en fonction de la pression P_0 . Les points importants sont repérés par des lettres. L'ordre alphabétique correspond à l'ordre chronologique des événements.

4.3.3 Variation de la longueur du résonateur aval

Les bifurcations qui traduisent les transitions entre les deux régimes d'oscillation sont également observées lors de modifications du couplage acoustique. La longueur du résonateur aval est modifiée en utilisant différents tubes en plexiglas, (cf. figure 4.1). L'augmentation de la longueur du tube a pour effet de diminuer la fréquence F_1 . Les géométries et impédances d'entrée estimées des résonateurs acoustiques sont présentées en annexe (figure D.1). Pour le tube en plexiglas le plus long ($L_p = 22 \text{ cm}$), la fréquence F_1 de la première résonance descend à 320 Hz .

Les tubes en plexiglas sont ajoutés et retirés manuellement au cours de la phonation. Pour chaque longueur L_p , la fréquence fondamentale moyenne F_{0Moy} des oscillations est déterminée sur la période durant laquelle le tube est en place. Ces périodes durent entre 4 et 8 s.

La figure 4.7 présente les résultats obtenus. La fréquence fondamentale F_0 est d'environ 219 Hz , lorsque L_p est inférieure ou égale à 10 cm , puis chute à 123 Hz pour des longueurs L_p supérieures ou égales à 12 cm . Ces valeurs de fréquence fondamentale sont caractéristiques des régimes RB et RA. Les figures 4.7.b et 4.7.c montrent que l'écart-type de F_0 augmente lorsque la valeur de L_p se rapproche de celle pour laquelle la bifurcation est observée. Les variations de F_{0Moy} sont plus importantes en régime A.

4.4 Simulation numérique

La validation des modèles théoriques est réalisée par le biais de simulations numériques visant à reproduire les mesures réalisées sur la maquette de l'appareil phonatoire.

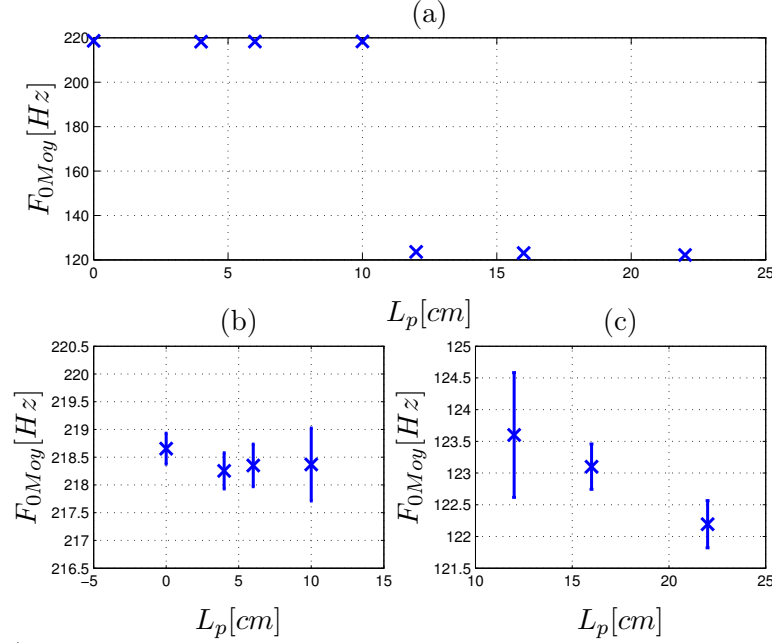


FIGURE 4.7 – (a) : Fréquence fondamentale moyenne F_{0Moy} des oscillations de la maquette de cordes vocales en fonction de la longueur L_p du tube en plexiglas, inclus dans le résonateur aval. Les valeurs sont moyennées sur 5 s de mesure. (b) : Agrandissement pour les oscillations appartenant au régime A, (c) : Agrandissement pour les oscillations appartenant au régime B. Les barres verticales représentent l'écart type associé à chaque mesure de F_0 .

4.4.1 Choix des paramètres du modèle à deux masses

Évolution des paramètres mécaniques avec la pression d'eau

La masse vibrante m est souvent estimée en considérant une partie de la masse d'eau totale contenue dans les répliques de corde vocale [Vil02, Rut07]. En utilisant des mesures de réponse mécanique effectuées sur les répliques de corde vocale, Rutty [Rut07] extrait les valeurs de k et de r à partir de la fréquence de la première résonance et de son facteur de qualité. Le choix des paramètres mécaniques repose sur les travaux de Hermant [Her14], réalisés sur la maquette de cordes vocales utilisée dans cette étude. Hermant utilise deux types de mesures pour estimer l'évolution des paramètres mécaniques en fonction de la pression d'eau P_e : des mesures de la pression d'alimentation P_{on} au seuil de l'oscillation et des mesures de la réponse mécanique des répliques de corde vocale. Les valeurs des paramètres mécaniques sont obtenues à partir des valeurs de P_{on} , des deux premières fréquences de résonance mécanique et des facteurs de qualité associés, en utilisant une méthode d'optimisation de type moindres carrés. La figure 4.8 présente les résultats obtenus.

Ces résultats montrent que la différence entre les valeurs des paramètres de masse, de raideur et d'amortissement associés aux masses 1 et 2 sont relativement faibles devant leurs ordres de grandeur. Ce constat justifie le choix suivant : $m_1 = m_2 = m$, $k_1 = k_2 = k$ et $r_1 = r_2 = r$, qui permet de réduire considérablement le nombre de paramètres de contrôle.

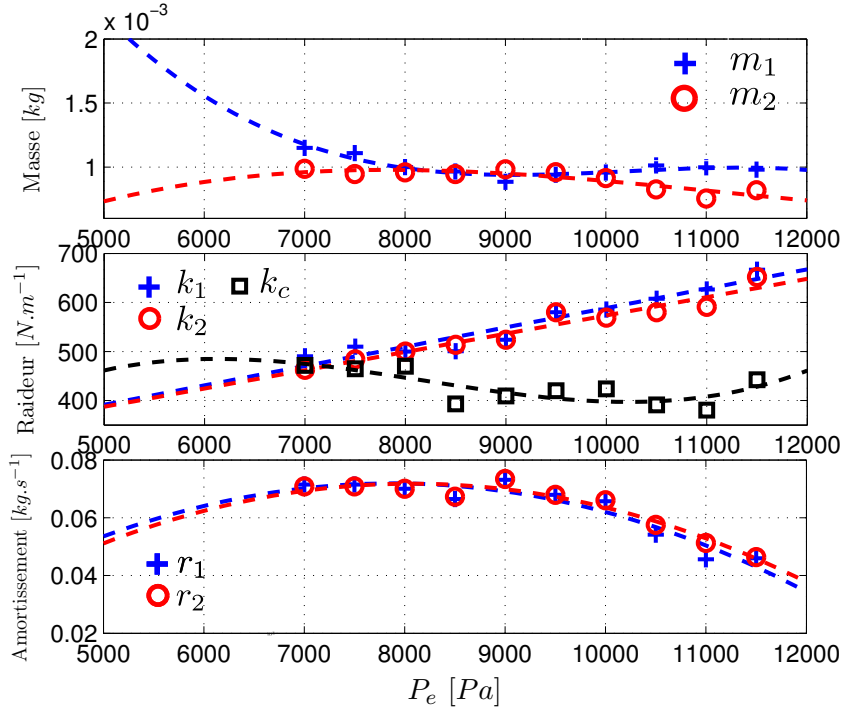


FIGURE 4.8 – Évolution des paramètres mécaniques du système à deux masses équivalent à la maquette de cordes vocales, en fonction de la pression d'eau P_e , d'après [Her14].

L'évolution des paramètres du système masse-ressort-amortissement équivalent en fonction de P_e peut être résumée de la façon suivante :

- la masse m diminue légèrement avec la pression P_e (environ $-5 \cdot 10^{-5} \text{ kg.Pa}^{-1}$),
- la raideur k augmente linéairement avec la pression P_e (environ $0,04 \text{ N.m}^{-1}.\text{Pa}^{-1}$),
- la valeur de la raideur de couplage k_c varie dans un intervalle défini par $0,6k \leq k_c \leq k$,
- le coefficient d'amortissement a tendance à diminuer avec la pression P_e (environ $-5 \cdot 10^{-6} \text{ s.m}^{-1}$), le même comportement est observé pour l'amortissement relatif ξ , défini par $\xi = r/(2\sqrt{mk})$ (environ $-3,3 \cdot 10^{-6} \text{ Pa}^{-1}$ avec $\xi \approx 0,025$ pour $P_e = 8000 \text{ Pa}$).

Dans la suite, l'hypothèse est faite que la pression P_e est une fonction affine de la raideur k telle que :

$$P_e = c_\alpha \times k + c_\beta, \text{ avec } c_\alpha = 20 \text{ Pa.m.N}^{-1} \text{ et } c_\beta = -1600 \text{ Pa.} \quad (4.1)$$

Les valeurs de c_α et de c_β sont choisies afin d'ajuster la fréquence fondamentale des simulations réalisées avec les données expérimentales. Ces valeurs ne correspondent pas à celles de la figure 4.8 (environ $c_\alpha = 30 \text{ Pa.m.N}^{-1}$ et $c_\beta = -7000 \text{ Pa}$). En effet, malgré le fait qu'il s'agisse de la même maquette de cordes vocales, les valeurs des paramètres mécaniques présentées en figure 4.8 ne peuvent pas être utilisées directement en raison des différences entre

Aérodynamique	$P_{sub} = 1300 \text{ Pa}$
Mécanique	$m = 0,34 \text{ g}, k = 480 \text{ N.m}^{-1},$ $\xi = 0,025, k_c/k = 0,8$ $\eta_k = 8 \cdot 10^6 \text{ m}^{-2}, \eta_r = 1 \cdot 10^7 \text{ m}^{-2}$ $\xi_c = 1, \alpha_{kc} = 0, \alpha_{\eta kc} = 0,$
Géométrie	$L_g = 16 \text{ mm}, \ell_g = 12 \text{ mm},$ $x_1 - x_0 = x_3 - x_2 = 1 \text{ mm}, y_0 = 0,05 \text{ mm}$ $h_0 = h_3 = 12,6 \text{ mm}$

TABLE 4.2 – Paramètres du modèle à deux masses correspondant au réglage de référence de la maquette de cordes vocales (cf. Tableau 4.1).

les réglages de la maquette. En particulier, la tension du boyau central en latex, très faible dans la présente étude, n'est pas quantifiée.

Choix des paramètres mécaniques et géométriques

La longueur $x_2 - x_1$ et la largeur L_g de la constriction glottique sont données respectivement par les diamètres des tubes en latex utilisés pour les répliques de corde vocale et pour le boyau central. Les valeurs des hauteurs h_0 et h_3 choisies correspondent au rapport de la section des résonateurs sur L_g . La valeur du paramètre η_k est ajustée afin d'obtenir, pour le régime A, une dépendance de F_0 en fonction de P_{sub} similaire à celle observée expérimentalement (cf. figure 4.12), égale $2,0 \cdot 10^{-2} \text{ Hz.Pa}^{-1}$. La valeur de η_k ainsi obtenue est du même ordre de grandeur que celle utilisée dans le modèle à deux masses de Ishikaza et Flanagan [IF72]. La valeur du paramètre η_r est celle proposée par Lucero [Luc05]. Les valeurs de k et de ξ pour $P_e = 8000 \text{ Pa}$ sont données par la figure 4.8. Les autres paramètres (m, k_c et y_0) sont fixés de manière à ajuster le mieux possible le résultat des simulations aux données expérimentales. Le tableau 4.2 présente les valeurs des paramètres utilisés pour les simulations visant à reproduire les mesures sur maquette.

4.4.2 Résultats

4.4.2.1 Variation de la raideur

Dans la simulation suivante, k est diminuée puis augmentée linéairement jusqu'à retrouver sa valeur initiale. La valeur minimum de k (275 N.m^{-1}) est atteinte à la moitié de la simulation, de durée totale 5 s. Durant cette variation de raideur, le rapport entre k et k_c est maintenu constant. De même, le coefficient d'amortissement r varie avec k de façon à ce que l'amortissement relatif ξ reste constant.

Le quotient ouvert Q_o des oscillations simulées est déterminé à partir des instants de fermeture t_f et d'ouverture t_o des cordes vocales (cf. équation 3.1). L'instant t_f correspond à l'instant de la première collision qui se produit au niveau de la première ou de la deuxième masse. L'instant t_o correspond à l'instant d'ouverture de la glotte, lorsque l'aperture h_i devient

positive au niveau des deux masses.

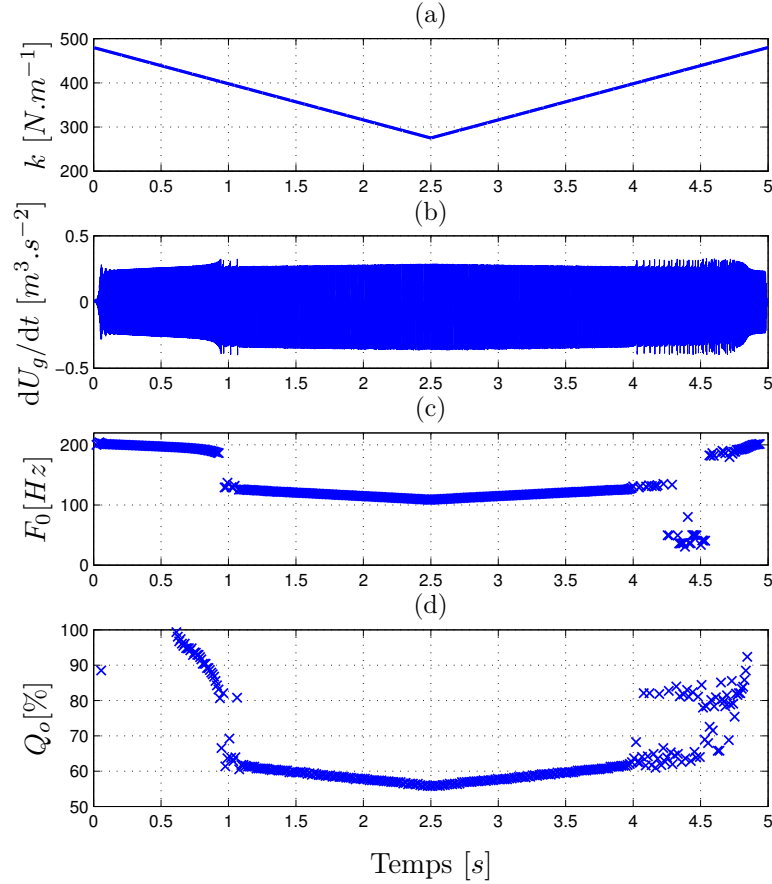


FIGURE 4.9 – Simulation d’une variation linéaire de la constante de raideur k , sans couplage acoustique. (a) : Constante de raideur k imposée. (b) : Dérivée temporelle dU_g/dt du débit glottique simulé U_g . (c) : Fréquence fondamentale F_0 du signal de débit glottique. (d) : Quotient ouvert Q_o , les valeurs ne peuvent être déterminées que lorsque la collision a lieu.

[Ecouter le signal \$dU_g/dt\$ \(son 4\)](#)

La figure 4.9 présente les signaux simulés dans le cas où le couplage acoustique n’est pas pris en compte. À $t = 1,1$ s, une discontinuité importante apparaît dans les signaux représentant dU_g/dt , F_0 et Q_o . Cet événement représente la transition entre un régime de fréquence fondamentale et quotient ouvert élevés ($F_0 > 170$ Hz, $Q_0 \geq 80$ %), appelé régime 2 (R2), et un autre régime pour lequel la fréquence et le quotient ouvert sont plus faibles ($F_0 \leq 140$ Hz, $Q_0 < 70$ %), appelé régime 1 (R1). Le régime 1 se caractérise par une fermeture complète de la glotte. À chaque période, la collision s’effectue au niveau des deux masses tandis que pour R2, la fermeture glottique, lorsqu’elle a lieu, s’effectue uniquement au niveau de la deuxième masse, celle située plus en aval de l’écoulement. La figure 4.10 illustre les motifs de h_1 et h_2 pour ces deux régimes.

Lors de l’augmentation de la raideur (figure 4.9), la transition R1→R2, est accompagnée d’un phénomène d’instabilité des oscillations, lorsque $4 < t < 4,7$ s. Durant cette phase, la

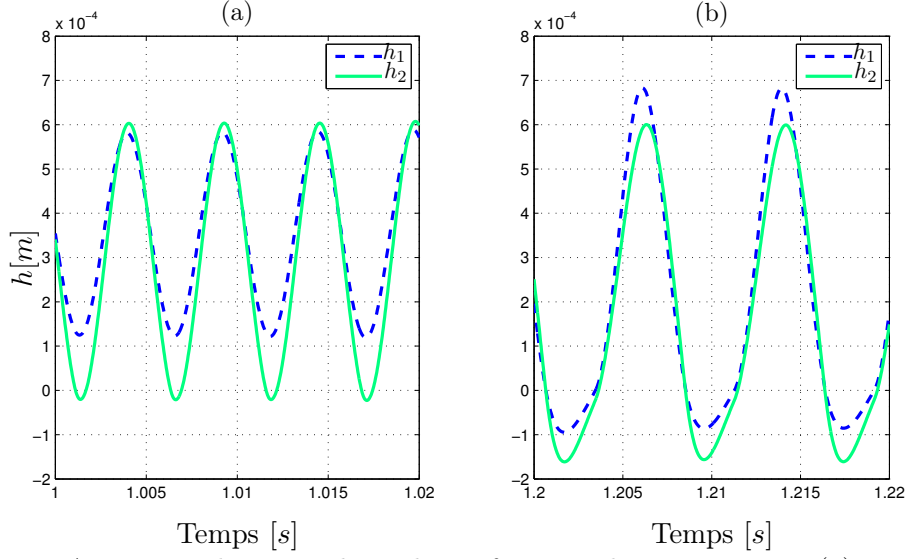


FIGURE 4.10 – Apertures glottiques h_1 et h_2 en fonction du temps, avant (a) et après (b) le premier saut de fréquence fondamentale de la simulation présentée en figure 4.9. Les signaux présentés par les graphiques (a) et (b) sont respectivement caractéristiques des régimes 2 et 1.

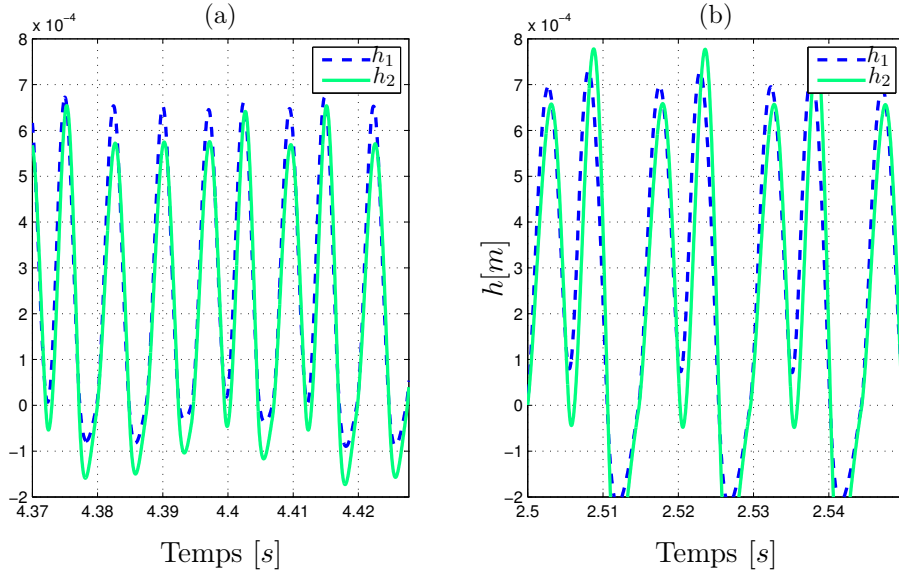


FIGURE 4.11 – Apertures glottiques h_1 et h_2 en fonction du temps, lors d'une transition R1→R2 (a) et en régime de doublement de période (b). Les signaux du graphique (a) correspondent à la simulation représentée en figure 4.9 et ceux du graphique (b) à ceux de la figure 4.12.b pour $P_e \approx 4100 \text{ Pa}$ et $F_0 = 68 \text{ Hz}$.

périodicité du signal dU_g/dt est plus faible et les valeurs de Q_o varient entre 60 et 85 %. Ce régime est caractérisé par l'organisation aléatoire des ses collisions, qui se produisent au niveau des deux masses ou uniquement au niveau de la première masse sans qu'il soit possible d'en extraire une structure de récurrence évidente (figure 4.11.a).

La figure 4.12 compare l'évolution des valeurs de F_0 mesurées et simulées dans le plan (P_e, F_0) . Lorsque le couplage acoustique est considéré (figure 4.12.b), l'effet d'hystérésis devient plus marqué. En effet, lors de l'augmentation de P_e , la transition R1→R2 ne se produit pas pour des valeurs de P_e inférieures à la valeur de référence ($P_e = 800 \text{ Pa}$).

Pour le régime 2, la fréquence fondamentale est plus élevée lorsque le couplage acoustique est pris en compte. Ce changement réduit l'écart relatif entre les amplitudes des sauts de fréquence mesurés et simulés, qui passe de 40 % à 26,7 % avec le couplage acoustique. Cependant, dans les deux cas, la valeur de F_0 correspondant à R1 est sur-estimée d'environ 8 %.

La figure 4.12.b montre que pour de faibles valeurs de la pression d'eau ($P_e < 4200 \text{ Pa}$), les simulations tenant compte de la propagation acoustique font apparaître un autre régime d'oscillation. Ce régime, appelé régime de doublement de période, se caractérise par des fréquences fondamentales estimées faibles ($F_0 \approx 68 \text{ Hz}$) et par le fait que les collisions entre les masses se produisent alternativement, comme celles observées en R1 et en R2 (cf. figure 4.11.b).

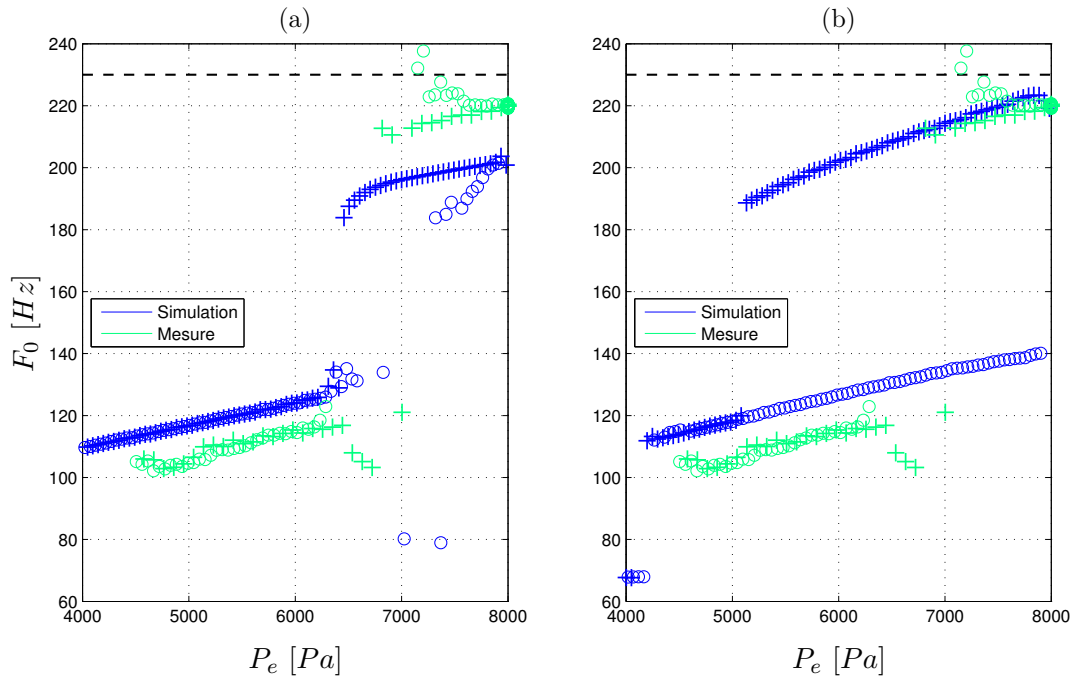


FIGURE 4.12 – Fréquence fondamentale F_0 en fonction de la pression d'eau P_e , sans (a) [Ecouler le signal \$dU_g/dt\$ \(son 4\)](#) et avec (b) [Ecouler le signal \$p_{av}\$ \(son 5\)](#) l'effet du couplage acoustique. Les symboles (+) et (o) représentent respectivement les trajectoires durant la diminution et l'augmentation de P_e . Pour la simulation, les valeurs de P_e sont déterminées à partir de k à l'aide de l'équation 4.1. Les traits pointillés représentent la fréquence de la première résonance du résonateur amont, estimée à partir de sa géométrie (cf. figure D.1).

4.4.2.2 Variation de la pression sous-glottique

Dans le but de reproduire les mesures présentées en section 4.3.2, des simulations sont réalisées avec une variation de la pression P_{sub} . À partir de la valeur de référence (1300 Pa), P_{sub} est augmentée jusqu'à 1700 Pa puis diminuée jusqu'à sa valeur initiale. Le signal de pression P_{sub} est construit de façon à empêcher les discontinuités dans la dérivée temporelle. La figure 4.13 présente les signaux obtenus.

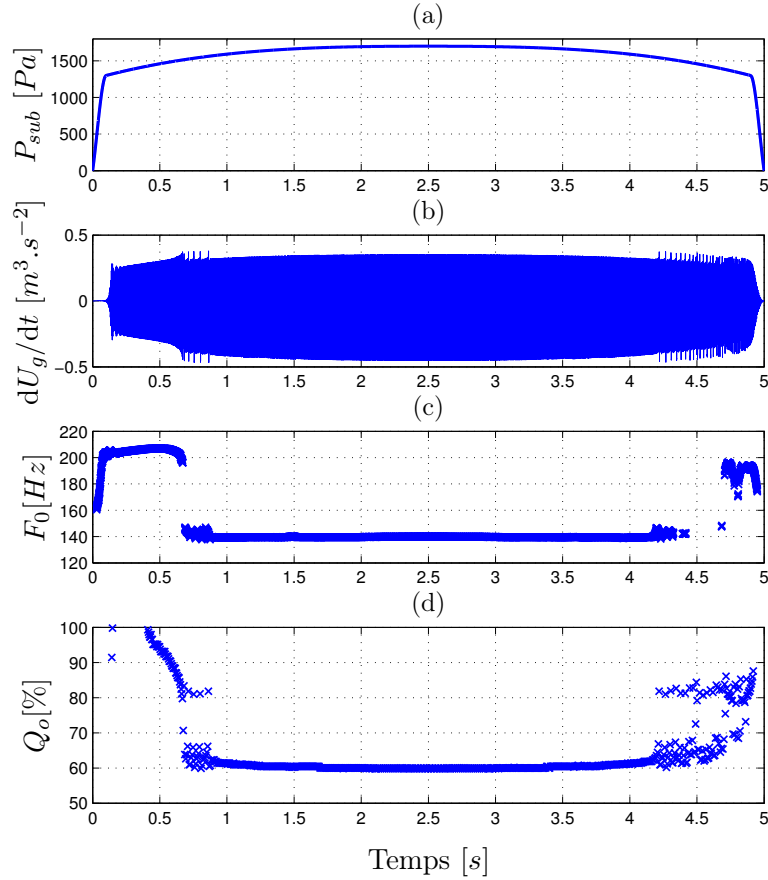


FIGURE 4.13 – Simulation d'une variation de pression sous-glottique P_{sub} . (a) : Pression sous-glottique P_{sub} imposée. (b) : Dérivée temporelle dU_g/dt du débit glottique simulé U_g . (c) : Fréquence fondamentale F_0 du signal dU_g/dt . (d) : Quotient ouvert Q_o , les valeurs ne peuvent être déterminées que lorsque la collision a lieu. [Ecouter le signal \$dU_g/dt\$ \(son 6\)](#)

Cette simulation montre qu'une augmentation progressive de P_{sub} provoque une bifurcation vers le régime 1. Cette transition se manifeste par la chute brutale des valeurs de F_0 et de Q_o à $t = 0,7$ s. Il est intéressant de noter que contrairement au cas précédent, F_0 augmente progressivement avec P_{sub} avant de chuter au moment de la transition. La transition dans le sens R1→R2 se produit lors de la diminution de P_{sub} , à $t = 4,8$ s. La grande variabilité du signal Q_o lors des transitions indique l'apparition de phénomène d'instabilité lorsque $0,7 < t < 0,9$ s et $4,2 < t < 4,8$ s.

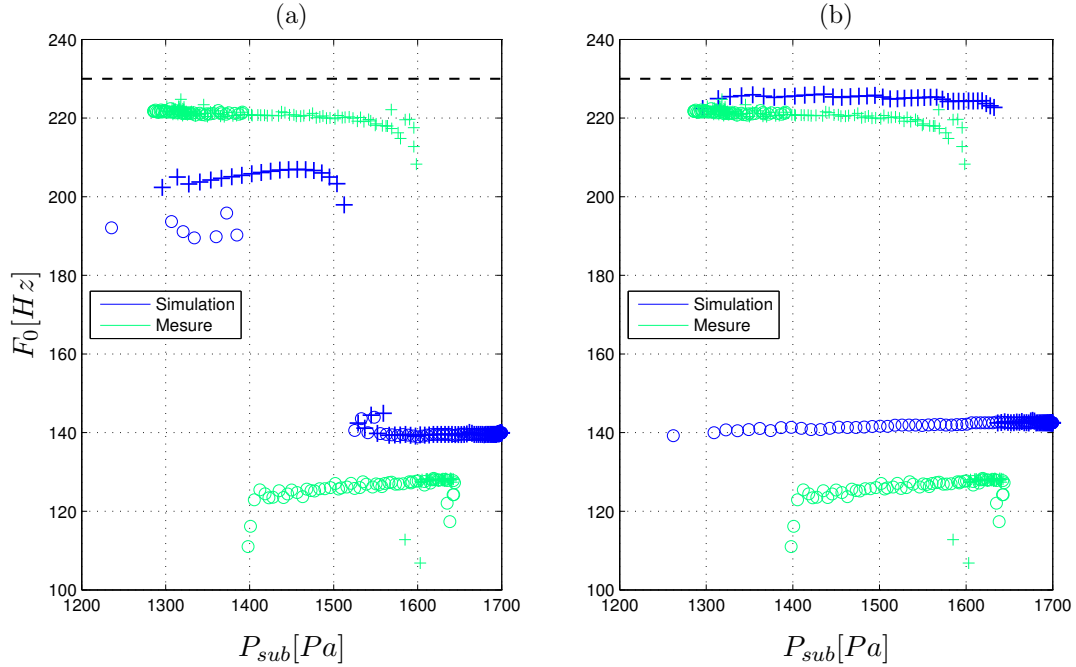


FIGURE 4.14 – Fréquence fondamentale F_0 en fonction de la pression sous-glottique P_{sub} , sans (a) [Ecouter le signal \$dU_g/dt\$ \(son 6\)](#) et avec (b) [Ecouter le signal \$p_{av}\$ \(son 7\)](#) l'effet du couplage acoustique. Le signal expérimental correspond à P_0 . Les symboles (+) et (o) représentent respectivement les trajectoires durant la diminution et l'augmentation de P_{sub} . Les traits pointillés représentent la fréquence de la première résonance du résonateur amont, estimée à partir de sa géométrie (cf. figure D.1).

La figure 4.14 compare les trajectoires simulées et mesurées du système dans le plan (P_{sub}, F_0) . À nouveau, l'hystérésis est sur-estimée dans le cas où le couplage acoustique est pris en compte : lors de la diminution de P_{sub} jusqu'à sa valeur initiale (1300 Pa), la transition R1→R2 ne se produit pas.

De manière analogue aux simulations de variation de la raideur, les prédictions sous-estiment l'amplitude du saut de fréquence. Lors de la transition R2→R1, les écarts relatifs sans et avec la prise en compte du couplage acoustique sont respectivement de 35 % et 7 %. Pour le régime 2, la prise en compte de l'interaction avec les résonateurs acoustiques modifie la dépendance de F_0 par rapport à P_{sub} . En estimant les trajectoires dans le plan (P_{sub}, F_0) par des droites, la pente estimée avec le couplage acoustique est du même ordre de grandeur que la valeur expérimentale ($-0,6 \cdot 10^{-2} \text{ Hz.Pa}^{-1}$), tandis que cette pente est positive dans le cas où le couplage acoustique n'est pas considéré (environ $2,4 \cdot 10^{-2} \text{ Hz.Pa}^{-1}$).

4.4.2.3 Variation de la longueur du résonateur aval

Les dernières simulations effectuées dans le cadre de la validation expérimentale sont réalisées pour une configuration statique, pour différentes longueurs de résonateur aval. Pour chaque simulation d'une durée 0,5 s, la fréquence fondamentale moyenne F_{0Moy} est extraite à partir des valeurs de F_0 lorsque $0,3 < t < 0,5$. La figure 4.15 présente le résultat de ces simulations.

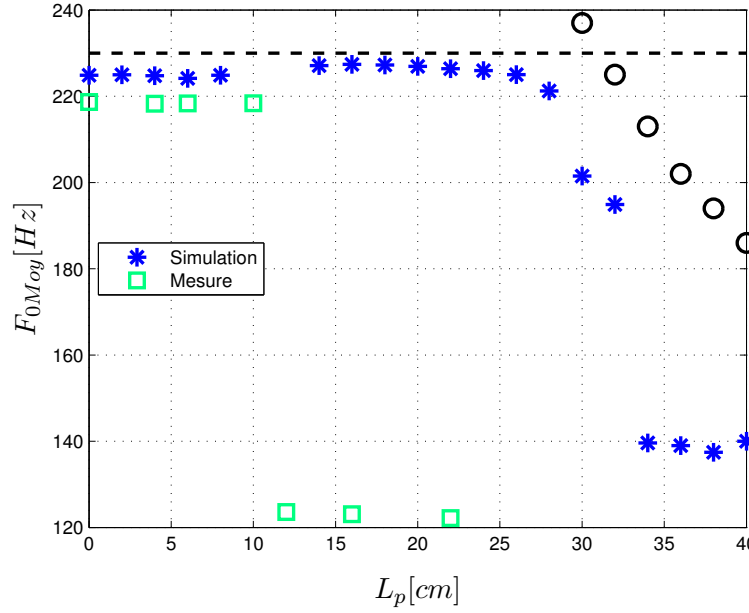


FIGURE 4.15 – Fréquence fondamentale moyenne F_{0Moy} en fonction de la longueur L_p du tube en plexiglas. Les symboles (---) et (o) représentent respectivement les fréquences de résonances des résonateurs amont et aval, estimées à partir de sa géométrie (cf. figure D.1).

Pour des longueurs $L_p \leq 32$ cm, les valeurs de fréquence fondamentale sont supérieures à $F_{0Moy} = 190$ Hz et inférieures à $F_{0Moy} = 145$ Hz, pour $L_p > 32$ cm. Cette séparation correspond à la frontière entre les régimes 1 et 2. En effet, la collision au niveau de la première masse ne se produit pas pour $L_p \leq 32$ cm.

Dans le cas où, $L_p = 32$ cm, la simulation fait apparaître un comportement qui présente des similitudes avec le régime présenté en figure 4.11.a à la différence près que, dans ce cas, il est possible d'identifier une structure récurrente. La collision au niveau de la première masse se produit une fois à chaque cycle de 10 périodes (cf. figure 4.16.a). Pour $L_p = 10$ cm et $L_p = 12$ cm, un phénomène de battement est observé sur les signaux de pression p_{am} et p_{av} , et d'ouvertures glottiques, h_1 et h_2 (cf. figure 4.16.b).

L'amplitude du saut de fréquence simulé lors de l'augmentation de L_p (55 Hz) est inférieure à la valeur observée expérimentalement (95 Hz), l'écart relatif associé est de 42 %. D'autre part, les simulations prédisaient que la transition R2→R1 se produit pour une longueur $L_p = 35$ cm, ce qui est supérieur à la valeur expérimentale ($L_p = 12$ cm).

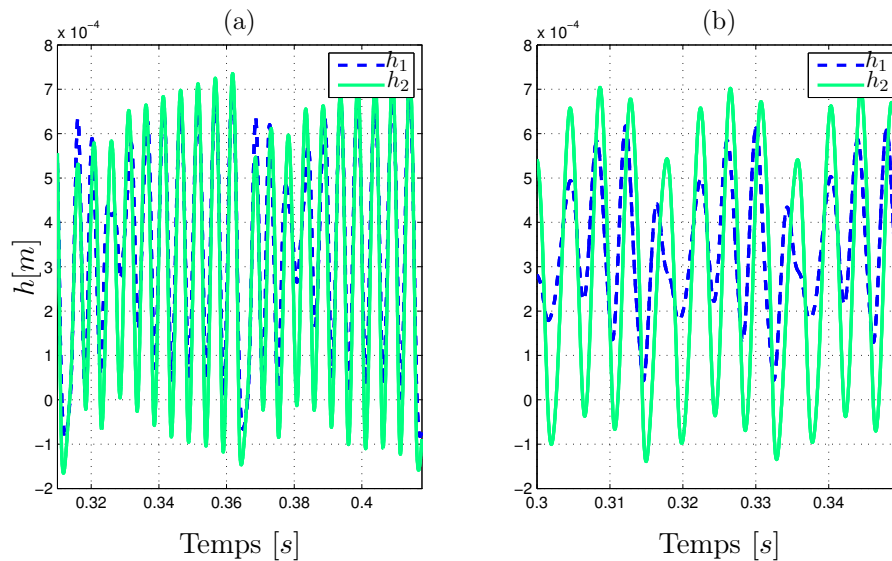


FIGURE 4.16 – Apertures glottiques h_1 et h_2 en fonction du temps, simulées pour $L_p = 32 \text{ cm}$ (a) et pour $L_p = 10 \text{ cm}$ (b).

4.5 Discussion et conclusion

4.5.1 Discussion

Mesures sur maquette

Les mesures mettent en évidence l'existence des deux régimes d'oscillations. Les transitions entre ces deux régimes se font sous la forme de bifurcations, la fréquence fondamentale de l'oscillation subit une brusque modification lors de la variation progressive d'un des paramètres de commande. La littérature ne mentionne pas l'observation de tels phénomènes sur une maquette de cordes vocales, dans le cas où les fréquences de résonance acoustique restent supérieures à celle de la fréquence fondamentale des oscillations.

Ces résultats peuvent cependant être mis en parallèle avec ceux obtenus sur des larynx excisés. En considérant que la pression d'eau dans les répliques de corde vocale est équivalente à la tension longitudinale, l'apparition d'un saut de fréquence négatif lors de la diminution de P_e est cohérente avec les résultats reportés dans de nombreuses études [dB60, BHTS96, SSM99, AFS09]. L'observation du même phénomène lors de l'augmentation de la pression sous-glottique est cohérente avec les observations faites par Berry et al. [BHTS96] et Alipour et al. [AFS09]. D'autre part, il semble que l'apparition d'une bifurcation vers un régime de fréquence fondamentale plus faible, lors de l'augmentation de la longueur d'un résonateur acoustique ne soit pas rapportée dans la littérature.

Dans certains cas, les transitions entre les deux régimes induisent un cycle d'hystérésis se traduisant par le fait que la trajectoire décrite par le système lors de l'augmentation du paramètre de contrôle est différente de celle décrite lors de la diminution. À l'approche de la transition la bifurcation est retardée, provoquant ainsi une plage de valeurs du paramètre de

contrôle pour lesquelles les deux régimes peuvent apparaître. Ce phénomène est particulièrement prononcé dans le cas où P_0 est utilisé comme paramètre de contrôle. Les transitions entre les deux régimes observées peuvent être accompagnées de phénomènes d'instabilité tels que le doublement de la période d'oscillation ou l'augmentation de la variance de la fréquence fondamentale.

Bien que les transitions entre les deux régimes aient également été observées en modifiant l'écartement entre les répliques de corde vocale, ce paramètre n'a pas été exploité mais uniquement utilisé comme paramètre de réglage, permettant ainsi d'ajuster la configuration de référence.

Dans cette étude, aucune mesure de la réponse mécanique des répliques de cordes vocales n'a été effectuée. Ces mesures fournissent des informations sur le comportement vibratoire de la structure [Rut07, Her14]. Les mesures effectuées par Hermant [Her14] donne, pour la même maquette de cordes vocales, une fréquence de la première résonance mécanique qui varie entre 80 et 120 Hz pour une pression d'eau comprise entre 4500 et 8000 Pa , c'est-à-dire des valeurs proche de la fréquence fondamentale de vibration mesurées dans cette étude.

Simulation numérique

Les simulations réalisées à l'aide des modèles théoriques présentés dans le chapitre 2 font apparaître deux régimes d'oscillation. Les transitions se manifestent sous la forme de bifurcations du système. L'analyse de l'évolution des apertures glottiques montre que la collision se produit au niveau des deux masses pour le régime 1 et uniquement au niveau de la deuxième masse en régime 2. Ainsi, la transition dans le sens R2→R1 est observée lorsque l'amplitude des oscillations devient assez importante pour provoquer l'apparition de collision au niveau des deux masses.

Les simulations réalisées montrent que les transitions sont parfois accompagnées par l'apparition de phénomène de doublement de période ou d'une organisation aléatoire des collisions successives. Les simulations font également apparaître un cycle d'hystérésis, amplifié lorsque le couplage acoustique est pris en compte. Ces phénomènes témoignent de la co-existence des deux régimes d'oscillation possible pour un même jeu de paramètres de contrôle. Ces résultats ne sont pas en accord avec ceux de Lucero [Luc96] qui observe, en utilisant le modèle à deux masses de Ishikaza [IF72], que la co-existence de plusieurs cycles limites ne se produit pas lorsque la formation du jet et les pertes visqueuses sont prises en compte dans l'écoulement. Pour les variations de k et de P_{sub} (figures 4.12 et 4.14), les résultats de simulation avec et sans prise en compte du couplage acoustique montrent que l'interaction acoustique avec les résonateurs n'est pas un phénomène physique moteur dans l'apparition de ces bifurcations mais modifie considérablement le système et influe notamment sur la valeur de la fréquence fondamentale de l'oscillation.

Les simulations les plus fidèles aux mesures sont obtenues sans aucune modification de la force élastique pour modéliser la collision entre les cordes vocales ($\alpha_{kc} = \alpha_{\eta kc} = 0$ dans l'équation 2.21). La première interprétation possible pour expliquer ceci, est que la collision entre les tubes en latex remplis d'eau n'est que partielle et que les effets de la collision sur les oscillations sont négligeables. La deuxième explication proposée est que modéliser la collision par une modification arbitraire du coefficient d'amortissement et de l'expression de la force

de rappel n'est pas adapté à une description correcte des chocs entre les répliques de corde vocale.

Validation expérimentale

Les oscillations en régimes 1 et 2 présentent des caractéristiques similaires à celles des mécanismes laryngés I et II observées chez l'humain. En effet, par rapport au régime 1, les oscillations en régime 2 se caractérisent par :

- une augmentation de la valeur du quotient ouvert,
- une réduction de l'amplitude des oscillations, en particulier au niveau de la première masse, qui représente l'extrémité amont de la glotte, induisant dans certains cas une absence totale de collision,
- une réduction de la surface de contact lors de la collision entre les cordes vocales.

Cependant, la validation expérimentale n'est valable que si les régimes A et B observés sur la maquette de cordes vocales sont induits par les mêmes phénomènes physiques que les mécanismes laryngés I et II et si les modèles théoriques utilisés pour les simulations sont capables de reproduire le comportement des régimes A et B.

Les régimes A et B observés sur la maquette de cordes vocales présentent certaines similitudes avec les mécanismes laryngés I et II. Les résultats obtenus en faisant varier la pression d'alimentation et la pression d'eau sont proches de ceux obtenus sur des larynx excisés, mentionnés dans la littérature (cf. section 3.2.1). Cependant, il est possible que pour les mesures présentées dans cette étude, l'interaction avec les résonateurs acoustiques joue un rôle important dans l'apparition du régime B. En effet, pour le régime B, les valeurs de fréquence fondamentale, comprises entre 210 et 230 Hz , sont proches de celle de la fréquence de la première résonance du résonateur amont, située à 230 Hz (cf. annexe D.1).

Pour les trois paramètres de contrôle étudiés, les simulations permettent de reproduire qualitativement l'évolution de la fréquence fondamentale des oscillations mesurée sur la maquette. D'un point de vue quantitatif, les prédictions de l'amplitude du saut de fréquence fondamentale lors de la transition entre deux régimes sont plus précises dans le cas où le couplage acoustique est considéré. En moyenne, les résultats de simulation sous-estiment cette valeur de 24 %. La prédiction des valeurs des paramètres de contrôle pour lesquels les bifurcations sont observées est moins fiable, en particulier dans le cas de la variation de la longueur du résonateur aval.

La capacité des modèles utilisés à reproduire les transitions de mécanismes laryngés peut donc raisonnablement être validée. Cependant, de par la simplicité de sa géométrie, le modèle à deux masses présente certaines limites. En effet, dans le modèle, la dimension de la constriction glottique dans la direction antero-postérieure est fixe alors que les observations *in vivo* rapportent que durant la phonation, les cordes vocales s'ouvrent et se ferment sur toutes leurs longueurs en mécanisme I et uniquement sur une partie de leur longueur totale en mécanisme II [Hen01]. De plus, il est évident que le comportement oscillatoire de structures continues telles que les cordes vocales et les répliques en latex remplies d'eau, ne peut être reproduit précisément par un système à deux degrés de liberté.

4.5.2 Conclusion

Bilan

Les mesures réalisées sur une maquette auto-oscillante de l'appareil phonatoire ont permis d'observer des sauts de fréquence fondamentale en faisant varier progressivement trois paramètres de contrôle différents. Ces sauts de fréquence marquent la transition entre deux régimes. Le régime A se caractérise par des valeurs de fréquence fondamentale comprises entre 100 et 125 Hz et le régime B par des valeurs comprises entre 210 et 230 Hz . Le régime A apparaît lors d'une diminution de la pression d'eau P_e dans les répliques de corde vocale, d'une augmentation de la pression d'alimentation P_0 en amont de la maquette ou bien lors d'une diminution de la fréquence F_1 , première résonance du résonateur aval. Les transitions prennent la forme de bifurcations et sont parfois accompagnées de phénomènes d'hystérésis, de doublement de la période et par l'apparition de régimes quasi-périodiques. Les bifurcations observées lors des variations de la pression de l'eau et de la pression d'alimentation sont cohérentes avec la littérature portant sur la phonation de larynx excisés.

La reproduction de ces expériences par simulation numérique fait également apparaître des sauts de fréquence lors de la transition entre deux régimes. Le régime 1 se caractérise par des valeurs de fréquence fondamentale comprises entre 110 et 140 Hz et de quotient ouvert inférieures à 70 % tandis que le régime 2 présente des valeurs de fréquence fondamentale comprises entre 180 et 220 Hz et de quotient ouvert supérieures à 80 %. Le motif de vibration des deux régimes observés diffère par la collision au niveau de la première des deux masses du modèle, celle-ci se produit uniquement en régime 1. Les transitions simulées présentent des caractéristiques similaires à celles observées sur la maquette tel que l'apparition de régimes quasi-périodiques.

Validation expérimentale

Les caractéristiques des régimes simulés 1 et 2 présentent certaines analogies avec celles des mécanismes laryngés observées sur des locuteurs humains. Pour le régime 2 (et respectivement le régime II), la surface et la durée du contact entre les cordes vocales opposées sont plus faibles que pour le régime 1 (et respectivement le régime I). De plus, le modèle à deux masses proposé par Lous [LHVH98] permet de reproduire qualitativement les transitions entre les régimes A et B. Lors d'une transition, l'amplitude du saut de fréquence observé sur la maquette est reproduite avec une précision de l'ordre de 25 %. La prédiction des valeurs du paramètre de contrôle à la transition est moins précise.

Les modèles théoriques présentés dans le chapitre 2 sont aptes à prédire l'apparition de bifurcations du système laryngé. Ces modèles seront utilisés dans le chapitre 5 pour étudier les différentes stratégies motrices qui conduisent à l'apparition d'une transition de mécanismes laryngés accompagné par un saut de fréquence fondamentale.

Perspective

Pour aller plus loin, il serait intéressant de reproduire ces mesures sur maquette en utili-

sant une caméra ultra-rapide afin de pouvoir étudier le mouvement oscillatoire des répliques de corde vocale lors des vibrations en régimes A et B. Cette information supplémentaire permettrait notamment de déterminer le quotient ouvert associé aux oscillations de la maquette.

Étude des stratégies motrices par simulation numérique temporelle

Sommaire

5.1	Introduction	93
5.2	Choix des paramètres de simulation	94
5.3	Méthode d'analyse des simulations	95
5.3.1	Détermination des paramètres acoustiques	95
5.3.2	Régimes d'oscillation	95
5.3.3	Méthode de classification des régimes d'oscillation	96
5.4	Résultats	97
5.4.1	Répartition des paramètres acoustiques	98
5.4.2	Répartition des régimes d'oscillation	103
5.4.3	Variation de la fréquence fondamentale	104
5.5	Discussion et conclusion	105
5.5.1	Discussion	105
5.5.2	Conclusion	107

5.1 Introduction

Les résultats obtenus au chapitre précédent montrent que le comportement vibratoire des cordes vocales peut se modéliser comme un système dynamique non linéaire et que les transitions de mécanismes laryngés, qui correspondent à des bifurcations du système, peuvent être décrites par les modèles théoriques simplifiés exposés au chapitre 2. L'objectif principal de ce chapitre est d'exploiter ces modèles afin de déterminer les stratégies motrices qui engendrent des bifurcations du système laryngé responsables d'un saut de fréquence fondamentale. L'influence du couplage acoustique sur ce type de bifurcations n'apparaît pas clairement dans la littérature. L'objectif secondaire de ce chapitre est d'éclaircir ce point. L'approche adoptée ici consiste à réaliser des simulations temporelles pour des valeurs différentes de paramètres de contrôle. Les résultats de simulation sont analysés en fonction des régimes d'oscillation observés et des paramètres acoustiques associés : la fréquence fondamentale F_0 et le quotient ouvert Q_o .

5.2 Choix des paramètres de simulation

L'ensemble des ajustements laryngés effectués par un locuteur pour produire une voyelle est réduit à trois paramètres du modèle physique. Le contrôle de la tension des cordes vocales est uniquement représenté par les variations de la raideur k . L'effort pulmonaire est décrit avec la pression sous-glottique P_{sub} . Enfin, le contrôle de la configuration géométrique de la glotte est réduit à la demi-aperture glottique y_0 qui représente le niveau d'adduction des cordes vocales.

Ces trois paramètres sont nommés paramètres de contrôle. Le tableau 5.1 présente les valeurs utilisées. Une simulation est réalisée pour chacune des 5832 combinaisons de valeurs de ces trois paramètres, pour deux valeurs de masses et pour quatre conditions de couplage acoustique différentes, ce qui correspond à un total de 46656 simulations.

Les autres paramètres du système à deux masses sont fixés à partir de valeurs utilisées par Lous et al. [LHVH98]. Le paramètre η_k représentant la contribution non linéaire de la force élastique est déterminé à partir de la valeur donnée par Ishizaka et Flanagan [IF72] et le paramètre η_r représentant la contribution non linéaire de la force visqueuse est choisi tel que $\eta_r = \eta_k$. Comme pour les simulations présentées au chapitre précédent (section 4.4), la collision est uniquement prise en compte au niveau de la force d'amortissement ($\alpha_{kc} = \alpha_{\eta kc} = 0$).

Paramètres de contrôle

	Minimum	Maximum	Pas
k [$N.m^{-1}$]	25	450	25
P_{sub} [Pa]	200	3600	200
y_0 [$10^{-5} m$]	2	36	2

Autres paramètres

Mécanique	$m = 0,1$ ou $0,05 g$ $\xi = 0,1$, $k_c/k = 0,6$ $\eta_k = 1 \cdot 10^6 m^{-2}$, $\eta_r = 1 \cdot 10^6 m^{-2}$ $\xi_c = 1,1$, $\alpha_{kc} = 0$, $\alpha_{\eta kc} = 0$,
Géométrie	$L_g = 20 mm$, $\ell_g = 2 mm$, $x_1 - x_0 = x_3 - x_2 = 0,2 mm$, $h_0 = 18 mm$, $h_3 = 22 mm$

TABLE 5.1 – Paramètres du modèle à deux masses utilisés pour la simulation. La partie supérieure du tableau présente les plages de variation des paramètres de contrôle. Ces variations sont linéaires : le pas entre deux valeurs successives est constant. La partie inférieure donne les valeurs constantes des autres paramètres du modèle à deux masses. Les simulations sont réalisées pour deux valeurs différentes de la masse m .

La durée de chaque simulation est fixée à $0,5 s$. La pression sous-glottique augmente progressivement pendant les 20 premières millisecondes de zéro jusqu'à sa valeur maximum (donnée par le tableau 5.1) puis diminue durant les 20 dernières millisecondes jusqu'à une

valeur nulle. Ces simulations sont réalisées pour deux valeurs de masse ($m = 0,1\text{ g}$ et $m = 0,05\text{ g}$) et pour quatre conditions différentes de couplage acoustique : sans couplage acoustique et pour les voyelles /a/, /i/ et /u/. Les résonateurs acoustiques aval et amont utilisés pour modéliser la trachée et le conduit vocal lors de la production de ces trois voyelles sont décrits en annexe (figure D.3).

5.3 Méthode d'analyse des simulations

Une fois les simulations réalisées, les données sont analysées en utilisant uniquement les signaux h_1 et h_2 correspondant aux apertures glottiques au niveau des deux masses.

5.3.1 Détermination des paramètres acoustiques

Pour chacune des simulations réalisées, l'évolution de la fréquence fondamentale du signal h_1 , au cours de la simulation, est estimée à l'aide de l'algorithme YIN [dCK02]. Le quotient ouvert est déterminé à partir des instants d'ouverture et de fermeture glottique. Les valeurs moyennes de ces deux paramètres acoustiques, Q_{oMoy} et F_{oMoy} , sont déterminées pour la durée T_s , pour laquelle $0.1 < t < 0.45\text{ s}$, de manière à ne pas prendre en compte les effets transitoires de début et fin de la simulation. Dans le cas des régimes chaotiques ou des régimes de doublement de période, ces valeurs moyennes n'ont que très peu de sens en raison de la valeur élevée de l'écart-type.

5.3.2 Régimes d'oscillation

Chaque réalisation simulée est classée automatiquement parmi quatre catégories de régimes d'oscillation.

Régime 1 (R1) [Ecouter un exemple de la pression \$p_M\$ simulée pour la voyelle /a/ \(son 8\)](#)

Le premier régime correspond à une oscillation stable pour laquelle les deux masses entrent en collision pour chaque période. Il est associé au mécanisme laryngé I.

Régime 2 (R2) [Ecouter un exemple de la pression \$p_M\$ simulée pour la voyelle /a/ \(son 9\)](#)

La seconde catégorie correspond à une oscillation stable pour laquelle la collision au niveau de la première masse ne se produit pas. Ce régime est associé au mécanisme laryngé II.

Régime C (RC) [Ecouter un exemple de la pression \$p_M\$ simulée pour la voyelle /a/ \(son 10\)](#)

Les réalisations associées à des oscillations auto-entretenues qui ne correspondent ni au régime 1 ni au régime 2, comme par exemple les régimes de doublement ou triplement de période ainsi que les régimes au comportement aléatoire, sont classées dans la troisième catégorie nommée régime C pour « chaotique ».

Régime E (RE)

La dernière catégorie correspond aux oscillations qui ne sont pas entretenues, dont l'amplitude décroît rapidement après l'augmentation de P_{sub} . Le déplacement des deux masses tend vers leurs positions d'équilibre.

5.3.3 Méthode de classification des régimes d'oscillation

5.3.3.1 Méthode ESPRIT

La classification des réalisations dans l'une des catégories repose sur l'analyse-synthèse du signal de l'aperture glottique h_1 au niveau de la première masse à l'aide de la méthode ESPRIT [RPK86, RK89]. Cette méthode d'analyse spectrale haute-résolution est basée sur la projection du signal sur un sous-espace « harmonique » et un sous-espace « bruit ». Cette méthode paramétrique suppose que la contribution harmonique du signal peut être décrite par une somme d'exponentielles complexes. Appliqué au signal h_1 , le modèle associé s'écrit comme la somme de la contribution harmonique \tilde{h}_1 et d'un bruit blanc b_r tel que :

$$\check{h}_1(t) = \tilde{h}_1(t) + b_r(t) = \sum_{k=1}^K a_k e^{j\phi_k} e^{\lambda_k t + j2\pi f_k t} + b_r(t). \quad (5.1)$$

Chaque exponentielle complexe d'indice k est décrite par 4 coefficients : l'amplitude a_k , la fréquence f_k , la phase ϕ_k et le facteur de croissance λ_k . La méthode ESPRIT consiste à estimer les paramètres des pôles λ_k et f_k pour un nombre de composantes K imposé par l'utilisateur. Les coefficients d'amplitude a_k et ϕ_k sont ensuite déterminés par la méthode des moindres carrés. Le signal estimé \tilde{h}_1 est obtenu à l'aide de l'équation 5.1.

5.3.3.2 Algorithme de classification

Bien que la méthode ESPRIT soit essentiellement utilisée pour analyser des signaux expérimentaux [Bad05, Sir13], notamment en raison de sa robustesse, les informations fournies par l'analyse-synthèse à l'aide de la méthode ESPRIT s'avèrent très utiles pour la classification des signaux simulés. En effet, ce processus présente un double intérêt pour distinguer les différents régimes d'oscillation.

D'une part, la différence entre les signaux simulés h_1 et reconstruits \tilde{h}_1 avec un nombre de composantes K relativement faible ($K = 5$) permet de dissocier les régimes chaotiques des autres types de régime. Le critère utilisé repose sur l'erreur quadratique relative entre les signaux \tilde{h}_1 et h_1 , définie par l'équation 5.2

$$E_{r1} = \left(\frac{\frac{1}{N_s} \sum_{n=1}^{N_s} (\tilde{h}_1[n] - h_1[n])^2}{\frac{1}{N_s} \sum_{n=1}^{N_s} h_1^2[n]} \right)^{1/2} \quad (5.2)$$

où N_s est le nombre d'échantillons temporels compris dans la durée T_s .

Dans le cas d'un régime chaotique, le motif complexe du signal h_1 est difficilement reproductible avec 5 composantes fréquentielles et l'erreur quadratique moyenne est importante, tandis que pour les régimes 1, 2 et E, le signal h_1 est reproduit plus précisément.

D'autre part, les valeurs des facteurs λ_k estimées permettent d'identifier les réalisations correspond au régime E. Lorsque les oscillations ne sont pas entretenues, l'amplitude des oscillations décroît rapidement au cours de la simulation et les valeurs de λ_k sont négatives.

L'analyse-synthèse est réalisée sur le signal h_1 simulé sur la période T_s . La figure 5.1 schématise l'algorithme de classification.

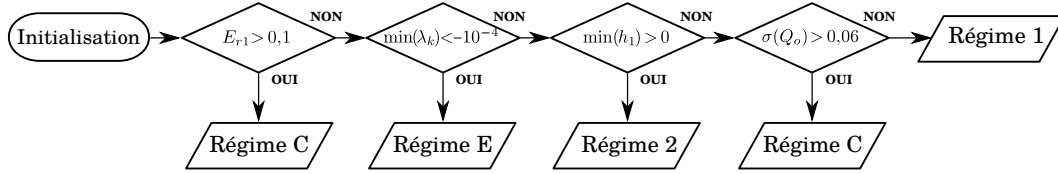


FIGURE 5.1 – Diagramme représentant l'algorithme utilisé pour assigner les réalisations numériques à l'une des quatre catégories de régime d'oscillation. La classification consiste en une série de tests représentés ici par des losanges. La notation $\sigma(Q_o)$ désigne l'écart-type du quotient ouvert Q_o .

Les tests effectués lors du processus de classification sont basés sur des seuils. Les valeurs des seuils sont déterminées par une démarche de type essai-erreur. La distinction entre les régimes 1 et C, déterminée par les critères sur l'erreur relative E_{r1} et sur la variance de Q_o , varie significativement en fonction des choix des valeurs de seuil. La démarche adoptée ici est de favoriser une certaine cohérence dans l'évolution des paramètres acoustiques F_0 et Q_o pour les régimes 1 et 2. Cependant, certaines erreurs de classification restent présentes en faible proportion. Il s'agit essentiellement de réalisations instables inexactement assignées au régime 1. Ces erreurs peuvent être repérées par leurs valeurs de quotient ouvert moyen Q_{oMoy} plus importantes que pour le reste des réalisations associées au régime 1.

5.4 Résultats

Les paramètres acoustiques associés aux régimes 1 et 2 sont présentés en figure 5.2. Les figures 5.3, 5.4 et 5.5 présentent respectivement les cartographies des régimes d'oscillation et des fréquences fondamentales moyennes simulées sans couplage acoustique, puis pour les

voyelles /a/ et /u/. Les résultats de simulation pour la voyelle /i/ et pour une masse plus importante sont présentés en annexe E.

5.4.1 Répartition des paramètres acoustiques

Fréquence fondamentale moyenne

La figure 5.2 montre que les variations de la fréquence fondamentale moyenne F_{0Moy} correspondant à R2 sont plus importantes (de 100 à 450 Hz) que celles associées à R1 (de 70 à 220 Hz). Les contours des régions associées aux régimes 1 et 2 permettent de déduire que pour un régime donné la fréquence augmente avec la raideur k et la pression P_{sub} , tandis qu'elle diminue légèrement avec y_0 .

Dans les cas où le couplage acoustique est faible (voyelle /a/) ou inexistant, l'évolution de la valeur de F_{0Moy} est comparable à celle de la fréquence de résonance mécanique F_m . Dans le cas où le couplage acoustique est fort (voyelle /u/), lorsque F_m se rapproche de la valeur de la première fréquence de résonance du conduit vocal $F_1 = 200$ Hz (c.f annexe D.2), les valeurs de F_{0Moy} associées à R1 et R2 se concentrent sur une plage de fréquence étroite ($150 < F_{0Moy} < 200$ Hz). Les valeurs de F_{0Moy} se rapprochent à nouveau de celles de F_m lorsque F_m atteint 400 Hz . La proportion de simulations pour laquelle la valeur de F_{0Moy} est comprise entre 200 et 400 Hz est très faible.

Quotient ouvert moyen

Le quotient ouvert moyen Q_{oMoy} est un critère pertinent pour distinguer les régimes 1 et 2. En effet, la figure 5.2 montre que sans couplage acoustique ou bien dans le cas de la voyelle /a/, les valeurs de Q_{oMoy} sont respectivement inférieures et supérieures à 80 % pour R1 et R2. Les quelques exceptions, pour lesquelles les valeurs de Q_{oMoy} associées à R1 sont supérieures à cette limite (sans couplage acoustique par exemple), sont dues à des cas de régimes chaotiques identifiés comme R1. La valeur de Q_{oMoy} ne descend pas en-dessous de 40 %.

La figure 5.2 permet d'analyser l'évolution globale du quotient ouvert moyen. Pour le régime 1, il apparaît que Q_{oMoy} augmente avec la raideur k et avec la demi-aperture glottique y_0 et diminue avec la pression sous-glottique P_{sub} . Pour le régime 2, les contours de la région associée ne permettent pas d'obtenir ces informations. Les simulations réalisées au chapitre 4 confirment que les mêmes comportements sont observés en R2 pour les variations de k et de P_{sub} (figures 4.9 et 4.13).

Dans le cas de la voyelle /u/, le renforcement du couplage acoustique fait apparaître, pour de grandes valeurs de P_{sub} et de y_0 , une augmentation des valeurs de Q_{oMoy} associées à R1, jusqu'à 90 %. Un recouvrement important apparaît alors entre les régions associées à R1 et R2.

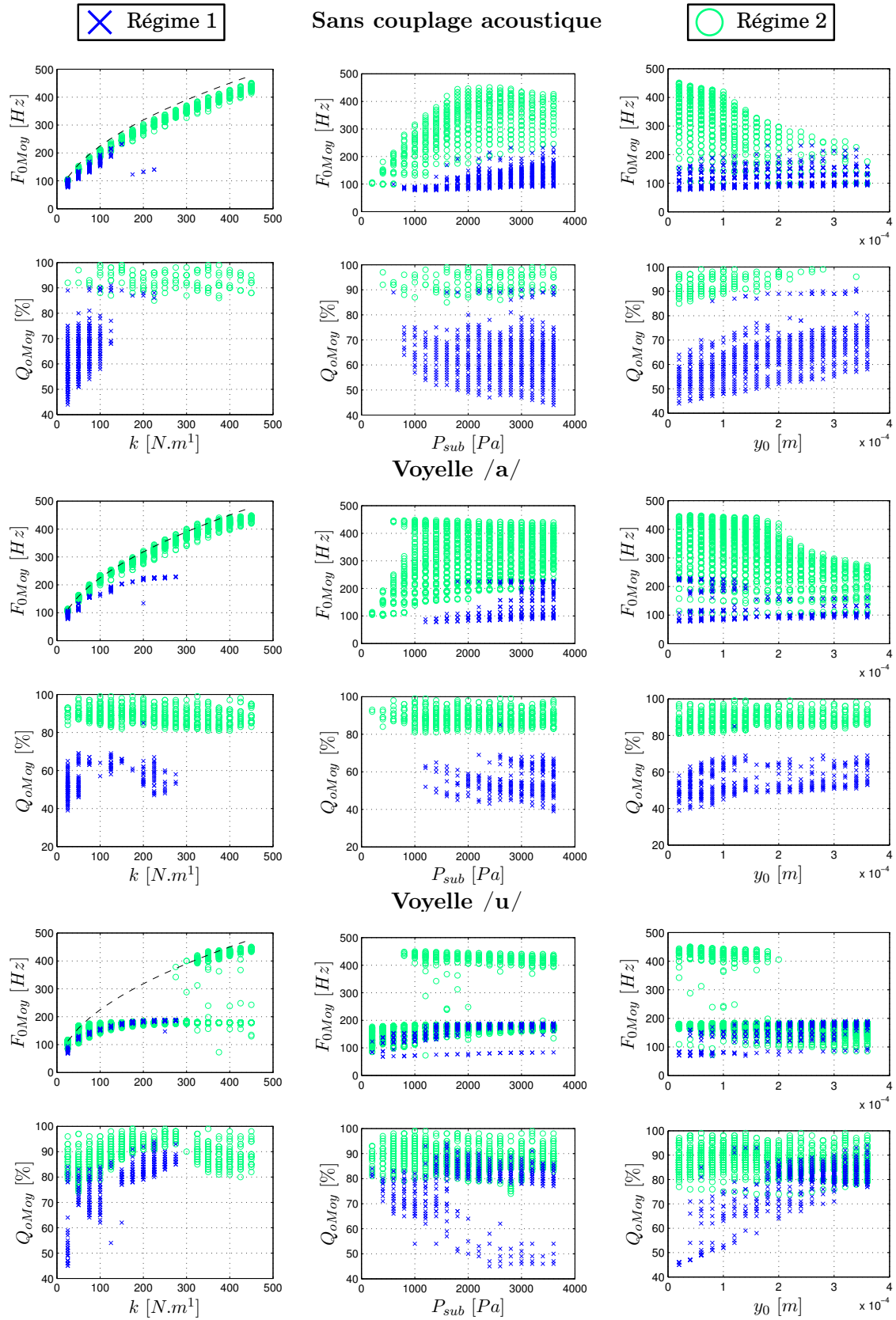


FIGURE 5.2 – Fréquences fondamentales F_{0Moy} et quotients ouverts moyens Q_{oMoy} avec $m = 0,05\text{ g}$, pour les régimes 1 et 2 en fonction de la raideur k , la pression sous-glottique P_{sub} et la demi-aperture glottique y_0 . Les courbes en pointillés représentent la fréquence de résonance mécanique « naturelle » des cordes vocales, définie par $F_m = \sqrt{k/m}/(2\pi)$.

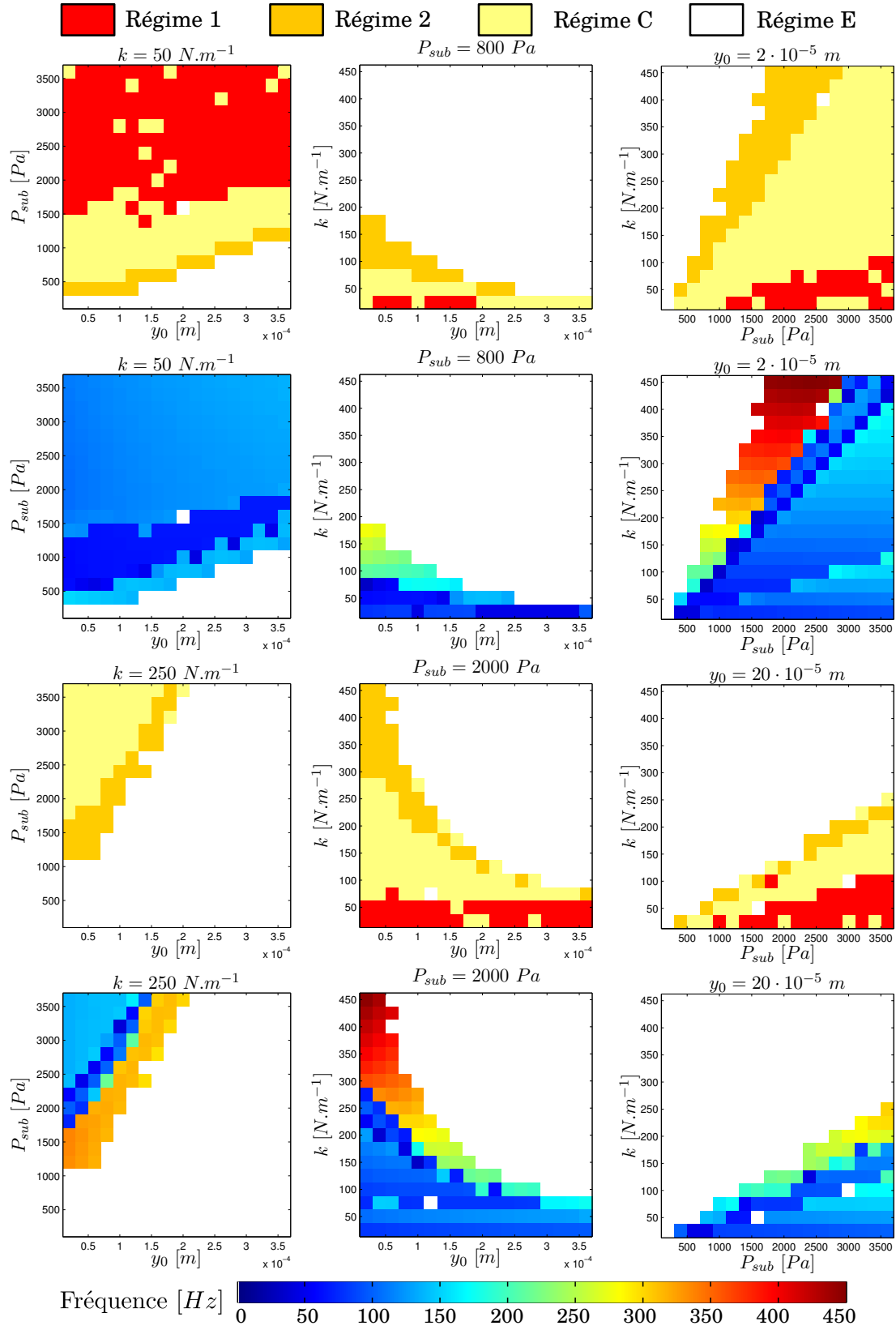


FIGURE 5.3 – Cartographie des régimes d’oscillation simulés sans couplage acoustique et pour une masse de $m = 0,05 \text{ g}$, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 .

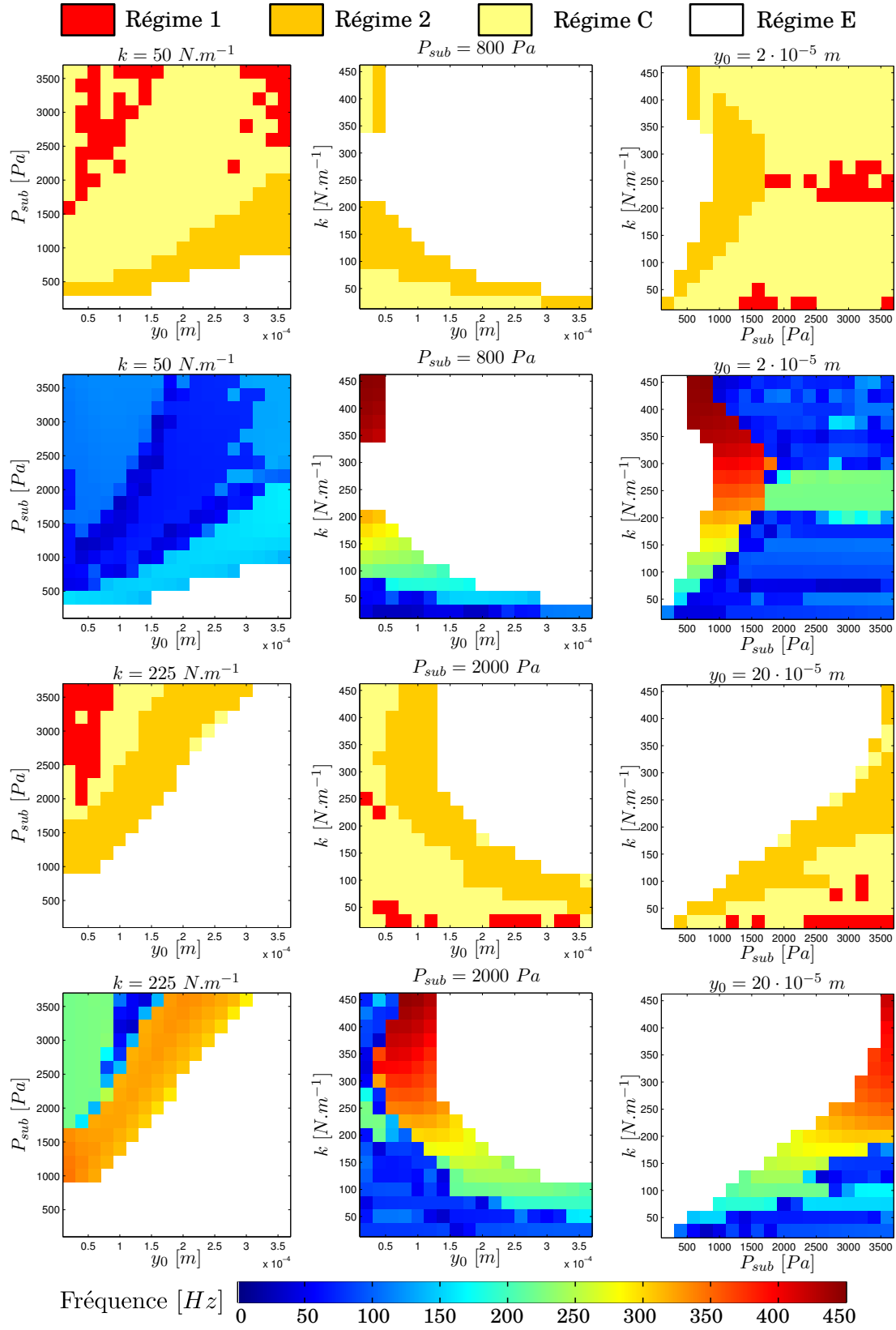


FIGURE 5.4 – Cartographie des régimes d'oscillation simulés pour la voyelle /a/ et pour une masse de $m = 0,05 \text{ g}$, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 .

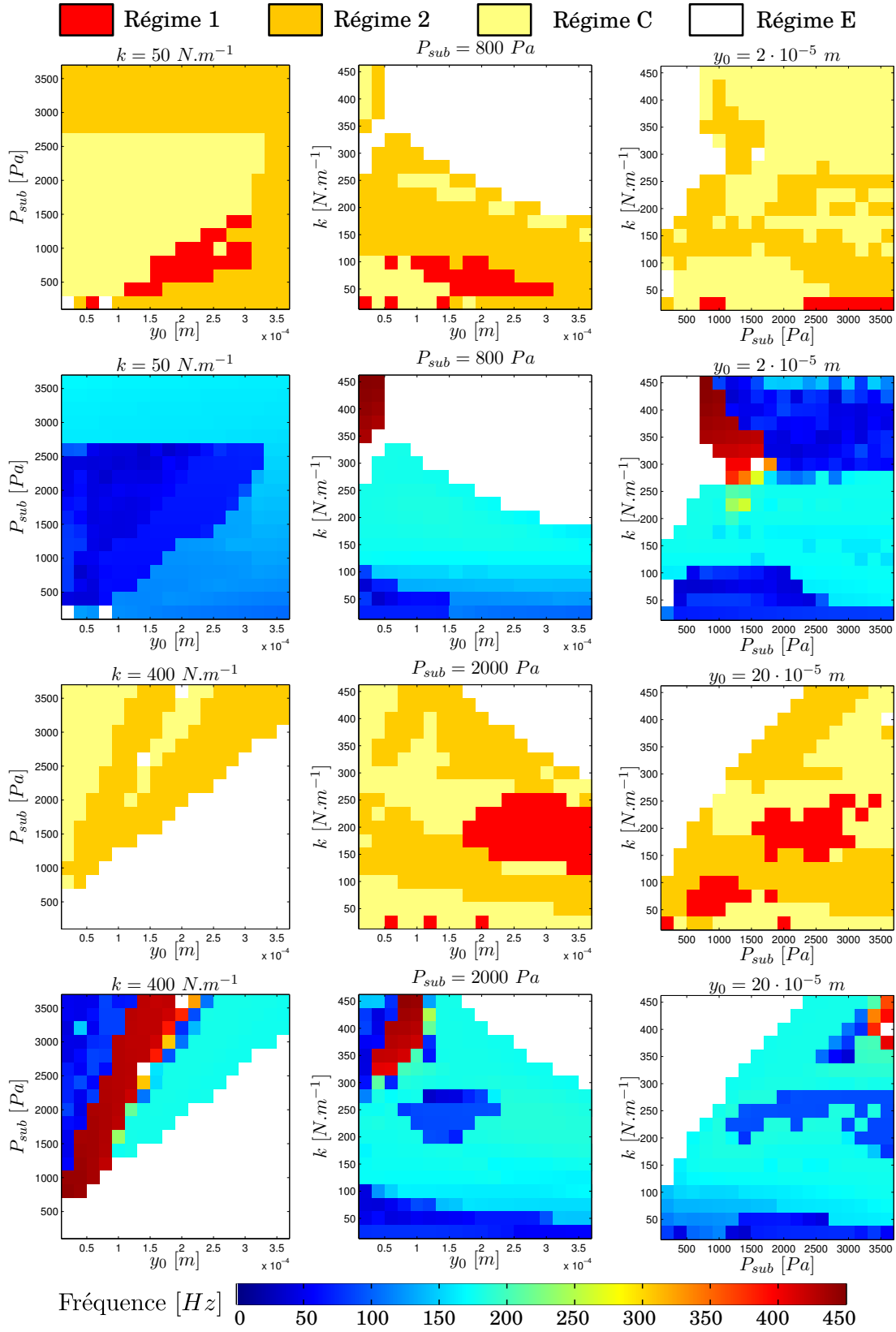


FIGURE 5.5 – Cartographie des régimes d'oscillation simulés pour la voyelle /u/ et pour une masse de $m = 0,05 g$, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 .

5.4.2 Répartition des régimes d'oscillation

Répartition globale

Le tableau 5.2 donne la proportion de chaque régime d'oscillation pour différentes configurations de résonateur acoustique et pour deux valeurs de la masse m . La répartition des régimes d'oscillation montre que les différentes conditions de couplage acoustique peuvent être séparées en deux catégories.

	$m = 0,1 \text{ g}$				$m = 0,05 \text{ g}$			
	s.ac.	/a/	/i/	/u/	s.ac.	/a/	/i/	/u/
Régime 1 [%]	8,7	2,3	22,4	22,8	9,3	5,0	16,3	11,4
Régime 2 [%]	8,4	17,0	31,0	37,6	7,9	19,4	26,0	37,0
Régime C [%]	21,1	22,7	30,6	25,4	20,3	23,3	34,1	28,5
Régime E [%]	61,9	58,0	16,0	13,7	62,6	52,3	23,5	23,1

TABLE 5.2 – Pourcentage de répartition de chaque régime d'oscillation pour les différentes voyelles et pour deux valeurs de la masse m . La notation s.ac désigne le cas où le couplage acoustique n'est pas pris en compte dans la simulation.

La première catégorie est celle pour laquelle le couplage est faible (voyelle /a/) ou inexistant. Pour les paramètres choisis, cette catégorie se caractérise par une proportion de régime E majoritaire, la proportion des régimes chaotiques (RC) est également supérieure à celle des régimes auto-entretenus et stables (R1 et R2).

La deuxième catégorie est celle pour laquelle le couplage acoustique est plus important en raison de la valeur plus faible de la première fréquence de résonance du conduit vocal (voyelles /i/ et /u/). Pour cette catégorie, la proportion de régimes auto-entretenus et stables augmentent au détriment du pourcentage de RE. Une légère augmentation de RC est également observée.

Sans couplage acoustique

Dans le cas où le couplage acoustique n'est pas pris en compte (figure 5.3), l'agencement des régions associées aux différents régimes est relativement organisé. Les régions associées aux régimes auto-entretenus sont systématiquement séparées du régime E par une zone frontière correspondant à R2. De même, les régions associées aux régimes 1 et 2 sont bien distinctes l'une de l'autre et elles sont séparées par une zone correspondant à des régimes chaotiques (RC).

Les deuxième et troisième colonnes de la figure 5.3 montrent qu'une augmentation de la raideur k a pour effet l'apparition de bifurcations successives dans le sens $R1 \rightarrow RC \rightarrow R2 \rightarrow RE$. Une augmentation de la pression sous-glottique P_{sub} (colonne 1 ou 3) induit des transitions de régime dans le sens inverse. Enfin, l'effet produit par une augmentation de la demi-aperture glottique y_0 est similaire à celui d'une augmentation de la raideur, bien que l'influence de y_0 soit moins importante pour les exemples présentés.

Voyelle /a/

Pour la voyelle /a/ (figure 5.4), l'organisation générale des régions associées aux différents régimes est semblable à celle décrite dans le paragraphe précédent. Cependant dans certains cas, l'apparition d'une deuxième région associée à R1 est observée. C'est par exemple le cas pour $k = 50 \text{ N.m}^{-1}$ (en haut à gauche), pour de grandes valeurs de P_{sub} et de y_0 ou bien pour $y_0 = 2 \cdot 10^{-5}$ (en haut à droite), lorsque $k \simeq 250 \text{ N.m}^{-1}$. La taille de la région associée à R2 est plus large que dans le cas où le couplage acoustique n'est pas pris en compte. De plus, il est observé que les régimes auto-entretenus gagnent du terrain sur RE (deuxième et troisième colonnes de la figure 5.4).

Voyelle /u/

Dans le cas où le couplage est plus fort, la figure 5.5 montre que l'organisation des régions associées aux différents régimes est considérablement modifiée. Le seul élément de l'organisation commun avec les cas précédents est que la région associée à RE est toujours séparée des régions associées aux régimes auto-entretenues par une zone frontière correspondant à R2. La répartition des zones associées à R2 est plus complexe et éclatée. Ainsi dans certains cas, par exemple pour $P_{sub} = 2000 \text{ Pa}$ et $y_0 = 20 \cdot 10^{-5} \text{ m}$, une augmentation de k peut par exemple provoquer des transitions successives de type $R2 \rightarrow R1 \rightarrow RC \rightarrow R2 \rightarrow RE$.

5.4.3 Variation de la fréquence fondamentale

Lorsque le couplage acoustique est faible ou inexistant, les deuxième et troisième colonnes des figures 5.3 et 5.4 montrent que, lors d'une transition dans le sens $R1 \rightarrow R2$ effectuée en augmentant uniquement la raideur k , l'augmentation de F_{0Moy} est plus importante lorsque l'épaisseur de la région associée au régime C à traverser est plus importante. Cette augmentation de F_{0Moy} est plus importante lorsque y_0 est petit et lorsque P_{sub} est grand.

Dans le cas où le couplage acoustique est plus important, les variations de F_{0Moy} sont plus limitées (cf. 5.4.1). Dans ces conditions, l'analyse de la figure 5.5 révèle que les variations de F_{0Moy} entre les deux régions associées à R1 et R2 sont très faibles, en particulier lorsque que la fréquence F_{0Moy} est proche de F_1 . Cependant dans certains cas, les variations de F_{0Moy} sont très importantes entre deux régions associées à R2. Pour $k = 400 \text{ N.m}^{-1}$ (en bas à droite), une diminution de la demi-aperture glottique y_0 provoque une augmentation de F_{0Moy} d'environ 200 Hz . Les exemples présentés en figure 5.5 montrent qu'une augmentation de la raideur peut produire le même effet.

5.5 Discussion et conclusion

5.5.1 Discussion

Stratégies motrices induisant une transition de mécanisme

Les résultats obtenus en section 5.4.2 montrent que les variations de chacun des trois paramètres testés conduisent à une transition de mécanisme laryngé. Une modification de la tension appliquée aux cordes vocales, qui peut être assimilée au paramètre de raideur k , est la stratégie la plus efficace pour provoquer une transition. Les simulations montrent également que l'influence de la pression sous-glottique P_{sub} est également importante. Ce résultat met en valeur le fait que les transitions de mécanisme ne peuvent pas être correctement expliquées par une analyse des valeurs et fréquences propres de la structure mécanique ne tenant pas compte de la pression sous-glottique [THvH07].

L'influence de chaque paramètre doit être considérée en fonction de l'effet produit sur l'amplitude des oscillations. En effet, une augmentation de l'amplitude des oscillations favorise l'apparition du phénomène de collision au niveau des deux masses, qui se traduit par une transition vers le régime 1. Les cartographies présentées en figure 5.3 et 5.4 montrent que lorsque l'effet du couplage acoustique n'est pas important, la région associée aux régimes « chaotiques » (régime C) se comporte comme une zone frontière entre les régions associées aux régimes 1 et 2. Cette observation peut s'interpréter par le fait que l'apparition de phénomènes d'instabilité ou de doublement de période témoigne de la coexistence de deux cycles limites associés à deux régimes différents.

Effet du couplage acoustique

Les simulations réalisées dans cette étude confirment que le couplage acoustique favorise l'apparition d'oscillations auto-entretenues [Rut07]. Dans le cas où la fréquence de résonance mécanique F_m est proche ou supérieure à la première fréquence de résonance du conduit vocal F_1 , ce couplage acoustique est fort et se traduit par un effet d'accrochage et de répulsion de la fréquence fondamentale F_0 des oscillations auto-entretenues. L'accrochage se produit à des fréquences F_0 inférieures aux fréquences de résonance du conduit vocal ou de la trachée, et la répulsion à des fréquences supérieures à celles des résonances acoustiques. Le fonctionnement de l'appareil phonatoire est alors semblable à celui d'un instrument à vent.

Dans le cas de la voyelle /a/, les figures 5.4 et 5.2 montrent que la simulation prédit également un effet d'accrochage pour la voyelle /a/ à une fréquence de 225 Hz, égale à la moitié de la valeur de la première fréquence de résonance de la trachée (cf. Annexe D.2).

L'effet d'accrochage est observé pour les régimes 1 et 2. Pour une valeur donnée de la raideur k , l'apparition de ce phénomène est favorisée lorsque la pression sous-glottique P_{sub} et la demi-aperture glottique initiale y_0 sont importantes (cf. respectivement figure 5.4 pour $y_0 = 2 \cdot 10^{-5} m$ et figure 5.5 pour $k = 400 N.m^{-1}$). Cet effet d'accrochage est réduit lorsque la masse m est plus importante (cf. figure E.4 pour la voyelle /a/).

Saut de fréquence fondamentale

L'analyse des simulations réalisées montre qu'il existe deux phénomènes induisant l'apparition de sauts de fréquence fondamentale. Le premier est la transition de mécanisme laryngé. Dans le cas où le couplage acoustique est faible ou inexistant, une augmentation de la raideur k induit une transition de régime dans le sens R1→R2 liée à une augmentation de la fréquence fondamentale. Cette augmentation est plus importante pour de grandes valeurs de P_{sub} et de faibles valeurs de y_0 . En interprétant de manière dynamique ces données obtenues en faisant varier les paramètres de contrôle de façon statique, ces résultats traduisent que l'intervalle du saut de fréquence fondamentale observé lors d'une transition du MI vers MII, est plus important lorsque la pression sous-glottique est grande et que le niveau d'adduction est important. Le deuxième phénomène qui explique l'apparition de sauts de fréquence est l'interaction acoustique entre les cordes vocales et les résonateurs situés en amont et en aval du larynx. Certains résultats obtenus pour la voyelle /u/ (figure 5.5 pour $k = 400 \text{ N.m}^{-1}$) suggèrent que les sauts de fréquence fondamentale induits par l'interaction acoustique peuvent s'accompagner de l'apparition des régimes d'oscillation identifiés comme instables.

Les résultats présentés ci-dessus montrent que lorsque le couplage acoustique est fort, les effets des interactions acoustiques sur l'auto-oscillation sont beaucoup plus importants que ceux liés aux régimes mécaniques. Dans ce cas, l'évolution de la fréquence fondamentale résulte essentiellement de l'effet d'accrochage sur les fréquences de résonance acoustique. Les transitions de régimes mécaniques n'ont alors plus d'influence sur la fréquence fondamentale (cf. section 5.4.3), ce qui peut s'interpréter par le fait que, dans ces conditions, les notions de régimes d'oscillation tels qu'ils sont définis en section 5.3.2 n'ont plus aucun sens physique.

Les simulations réalisées montrent qu'en fonction des conditions de couplage acoustique, la même stratégie motrice peut induire des effets opposés. Par exemple, dans le cas de la voyelle /a/ (figure 5.4 pour $k = 225 \text{ N.m}^{-1}$), une augmentation de y_0 produit un saut de fréquence fondamentale positif, associé à une transition de R1 vers R2. Dans le cas de la voyelle /u/, une augmentation de y_0 , qui a pour effet d'augmenter l'influence des interactions acoustiques et de diminuer légèrement F_0 , provoque un accrochage vers une fréquence de résonance acoustique inférieure, sans transition de régime. Il en résulte un saut de fréquence négatif.

Au vu de ces constatations, il apparaît difficile de résumer l'effet du couplage acoustique sur l'amplitude du saut de fréquence et sur la fréquence à laquelle se produit la transition de régime.

Limites de la simulation

Les simulations effectuées dans ce chapitre prédisent que l'influence du couplage acoustique est très importante en particulier pour les voyelles fermées. Dans les cas où se produisent les phénomènes d'accrochage de F_0 sur les fréquences de résonance du conduit vocal ou de la trachée, l'analyse détaillée des simulations montre que l'amplitude de la pression acoustique en amont ou en aval du larynx est considérable. Dans le cas de la voyelle /u/, pour les nombreuses simulations pour lesquelles F_0 est comprise entre 160 et 200 Hz, par exemple pour $k = 50 \text{ N.m}^{-1}$ et $P_{sub} > 2700 \text{ Pa}$ (figure 5.5), l'amplitude crête de la pression p_{av} atteint la valeur de P_{sub} . Le même phénomène est observé pour la pression acoustique amont p_{am} .

lorsque F_0 s'approche de la première fréquence de résonance de la trachée, par exemple pour $P_{sub} = 800 \text{ Pa}$ (figure 5.5), pour des valeurs de raideur k élevées et d'aperture glottique y_0 faibles.

Ces observations suggèrent que les simulations sur-estiment l'influence des interactions acoustiques entre les résonateurs et les cordes vocales. L'origine de ce phénomène peut être liée au couplage acoustique, plus particulièrement à l'approximation d'onde plane au niveau de la glotte [LM05, SA15]. Une autre explication possible est que le modèle de propagation acoustique sur-estime l'amplitude de l'impédance d'entrée du conduit vocal au niveau des résonances. Les mesures estiment que l'amplitude du premier pic est de l'ordre de 10^6 Pa.s.m^{-3} [HSW12], tandis que le modèle acoustique utilisé donne, pour le premier pic, une amplitude de l'ordre de 10^7 Pa.s.m^{-3} (figure D.3). Cette différence est possiblement due fait que, dans le modèle utilisé, les parois du conduit vocal sont considérées comme parfaitement rigides. En réalité, l'impédance de ces parois n'est pas infinie et elles contribuent à absorber de l'énergie acoustique et à la transférer vers l'extérieur.

5.5.2 Conclusion

Bilan

Des simulations ont été réalisées en faisant varier la raideur des cordes vocales, la pression sous-glottique et la demi-aperture glottique initiale, pour deux valeurs de masse différentes et pour différentes conditions de couplage acoustique. Les simulations ont été analysées en fonction du régime d'oscillation, de la fréquence fondamentale et du quotient ouvert. L'interprétation des résultats montre que les sauts de fréquence fondamentale peuvent être provoqués par deux phénomènes différents. Le premier est la transition entre deux régimes d'oscillation, dont l'apparition peut être induite par de nombreuses stratégies motrices différentes. Le deuxième est le phénomène d'accrochage de la fréquence fondamentale sur les fréquences de résonance acoustique. Ce phénomène est caractéristique des instruments à vent pour lesquels le couplage est fort et la fréquence fondamentale quasiment imposée par la géométrie de l'instrument. Ces deux phénomènes peuvent potentiellement interagir ensemble. Les effets d'une stratégie motrice donnée dépendent donc des conditions de couplage acoustique.

Les simulations permettent de reproduire des transitions de régimes lorsque le couplage acoustique n'est pas pris en compte. L'origine de ce phénomène n'est donc pas acoustique, bien que les interactions entre les deux phénomènes font que, dans certains cas, une modification de la géométrie des résonateurs entraîne une transition de régime.

Hypothèses sur la stratégie motrice employée par le sujet Y

Plusieurs hypothèses sont proposées pour expliquer les transitions de mécanisme laryngé produites par le sujet sourd, présentées dans le chapitre 3 et dans l'annexe C. Les explications proposées partent du principe que ces transitions dans le sens MII→MI sont accompagnées par un mouvement d'abaissement de la langue (section cf. 3.3.3).

- ① La première explication possible est basée sur la fréquence fondamentale intrinsèque des voyelles (cf. section 1.3.2). L'élévation de la langue, nécessaire à la production de voyelle fermées, entraînerait une tension des cordes vocales importantes, ce qui se traduirait par une production initiale en mécanisme II. À la fin de la production, le relâchement de la position haute de la langue entraînerait une diminution de la tension des cordes vocales et donc l'apparition du mécanisme I.
- ② La deuxième explication de l'origine de la transition MII→MI repose sur l'interaction acoustique avec le conduit vocal. L'abaissement de la langue a pour effet d'augmenter la valeur de la fréquence de la première résonance F_1 . En supposant qu'avant le saut de fréquence, F_1 soit inférieure à F_0 , l'abaissement de la langue peut amener à la situation où F_1 devient supérieure à F_0 . Cette condition de couplage favorise les oscillations dont l'amplitude augmenterait jusqu'à l'apparition du mécanisme I. Cependant cette hypothèse n'explique pas pourquoi la production est initiée en mécanisme II, avec une valeur de fréquence fondamentale si importante.

Il est également possible que ces deux processus contribuent tous les deux à l'apparition de la transition de mécanisme laryngé. L'hypothèse selon laquelle la transition de mécanisme observée dans le sens MII→MI est provoquée par une augmentation de la pression sous-glottique est peu probable pour plusieurs raisons. Une diminution de F_0 est observée dans les instants qui précèdent la transition, or une augmentation de la pression sous-glottique produit l'effet inverse. De plus, un abaissement de la langue a pour effet de diminuer la résistance au passage de l'air dans l'ensemble du système phonatoire, ce qui pour un effort pulmonaire constant, engendre une diminution de la pression sous-glottique. Enfin, le fait que pour le sujet Y, ces transitions soit observées en fin de production suggère que l'effort pulmonaire diminue progressivement à ce moment. La transition observée résulterait donc d'un problème de synchronisation. À la fin de la production, la position des articulateurs est relâchée alors que le niveau de pression sous-glottique est maintenu pendant environ 100 ms à un niveau suffisamment important pour entretenir la mécanisme I.

Perspectives

Les simulations montrent que les modèles théoriques utilisés sur-estiment l'effet du couplage acoustique. Ce point pourrait être amélioré en suivant les propositions de Laje et Mindlin [LM05], qui considèrent une distance de relaxation entre la glotte et le point du conduit vocal à partir de laquelle l'approximation d'onde plane est valide. Selon Sciamarella et Artana [SA15], ce raffinement dans la modélisation du couplage acoustique réduit considérablement l'effet du couplage acoustique. Dans le cas où les résonateurs acoustiques sont constitués de tissus humains, la modélisation de la vibration des parois du conduit vocal et de la trachée pourrait également réduire l'influence du couplage acoustique. Dans cette optique, Story [Sto95] propose un formalisme adapté au formalisme de synthèse par guide d'onde, qui consiste en l'ajout d'un système dynamique linéaire du second ordre au niveau de chaque jonction entre les tronçons cylindriques.

Troisième partie

Séquences voyelle - plosive non voisée - voyelle

Description de la troisième partie

La dernière partie de ce manuscrit porte sur la modélisation physique de la production de séquences voyelle - plosive non voisée - voyelle. Les modèles théoriques proposés et évalués à l'aide de mesures réalisées sur la maquette de l'appareil phonatoire, décrivent uniquement les interactions entre les cordes vocales et le conduit supra-laryngé. En effet, le bruit aéroacoustique qui caractérise la production des consonnes plosives lors du relâchement de l'occlusion n'est pas modélisé. Pour des raisons pratiques, les expériences et les simulations réalisées sont appliquées au cas des plosives bilabiales, les modèles développés sont cependant valables dans les cas où l'occlusion du conduit vocal est réalisée en amont des lèvres.

Le chapitre 6 traite de la construction d'un modèle aérodynamique permettant de prédire les effets de la réalisation d'une occlusion du conduit vocal sur les oscillations des cordes vocales. Le point de départ de ces travaux est le modèle à deux masses proposé par Lous et al. [LHVH98]. Le travail réalisé porte sur l'écoulement dans les lèvres et sur l'expansion passive de la cavité supra-glottique.

L'étude présentée au chapitre 7 utilise le modèle d'écoulement dans les lèvres proposé au chapitre 6. Les effets de la propagation et du couplage acoustique sur les oscillations sont pris en compte afin d'étudier leur impact sur le délai d'établissement du voisement après le relâchement de l'occlusion (*Voice Onset Time*).

Modélisation de l'interaction aérodynamique entre les cordes vocales et le conduit vocal

Sommaire

6.1	Introduction	114
6.2	Measurements on a human speaker	115
6.2.1	Example of an /apa/ sequence	115
6.2.2	Influence of the cheeks expansion	117
6.3	Air flow model for the lip constriction	118
6.3.1	Theoretical models	118
6.3.2	Experimental validation on a replica of the phonatory system	119
6.4	Aerodynamic model of the supraglottal cavity expansion	123
6.4.1	Theoretical models	123
6.4.2	Experimental validation by numerical simulations	126
6.5	Numerical simulations of an /apa/ sequence	128
6.5.1	Parameters of the simulations	128
6.5.2	Results	129
6.6	Conclusion	130

Note Liminaire

Ce chapitre, rédigé en anglais, est la reproduction d'un article intitulé « *Modeling of aerodynamic interaction between vocal folds and vocal tract during the production of vowel-voiceless plosive-vowel sequence* », soumis le 7 Avril 2015 dans la revue **Journal of the Acoustical Society of America**.

Résumé en Français

Le contexte de cette étude est la modélisation physique de la production de parole. L'objectif est de construire un modèle aérodynamique approprié pour étudier l'interaction entre les cordes vocales et le conduit vocal, lors de la production d'une séquence voyelle - plosive

non voisée - voyelle. La première étape de ce travail est de réaliser des mesures acoustiques et aérodynamiques sur un locuteur lors de la production de la séquence /apa/. L'aperture et la largeur des lèvres sont également mesurées à partir de l'enregistrement des mouvements du visage du sujet à l'aide d'une caméra ultra-rapide. Dans un second temps, des modèles théoriques décrivant l'écoulement d'air à travers les lèvres et l'effet d'une expansion de la cavité supraglottique sont proposés et puis validés à partir des mesures effectuées sur la maquette du système phonatoire. Enfin, les modèles sont utilisés pour réaliser des simulations numériques de la séquence /apa/, en utilisant les dimensions labiales mesurées comme uniques paramètres de contrôle variant dans le temps. Les simulations indiquent que la réalisation d'une occlusion du conduit vocal produit une augmentation passive de l'aire glottique associée à l'extinction de la phonation et que l'expansion de la cavité supraglottique est responsable du maintien de la phonation jusqu'à 40 ms après la fermeture des lèvres.

Remarques

Dans ce chapitre, le modèle théorique utilisé pour décrire le comportement mécanique des cordes vocales est celui employé par Lous et al. [LHVH98] et donc ne correspondent pas exactement à celui présenté dans le chapitre 2 de ce manuscrit. En effet, les expressions des forces élastique et visqueuse dans le modèle mécanique proposé par Lous et al. sont linéaires et les modifications des coefficients de raideur et d'amortissement lors de la collision sont différentes de celles proposées dans le chapitre 2. D'autre part, la maquette de cordes vocales et le dispositif expérimental utilisé dans ces travaux correspondent à ceux décrits au chapitre 4.

6.1 Introduction

The production of a vowel - plosive - vowel sequence involves several successive steps. First, during the initial vowel, the vocal folds self-oscillate. Then, by a voluntary articulatory movement, a complete occlusion of a portion of the vocal tract (lips, tongue, epiglottis or glottis) is realized and held. The vocal folds then stop to oscillate until the occlusion is released while the plosive sound is produced. Lastly, during the final vowel, the vocal folds start to vibrate again. In the case when a voiced plosive is involved, the vocal folds continue to vibrate during the closure of the vocal tract. Numerous studies on these productions have been carried out using acoustic measurements [LDC57, LA64, SK74, Kla75, Zue76, Byr93] or articulatory measurements [Per69, Wes83, BGBN87, LG97], while still very little attempts to model the production have been proposed.

One popular model for plosive production is proposed by Rothenberg [Rot68]. This lumped elements theoretical model is based on a low-frequency modeling of the aerodynamics in the glottis and in the vocal tract. Following this approach, numerous analysis of data from measurements on speakers [MW80, Wes83, Kea84, WK86, Dar87] have been carried out. All of these works take into account for an air flow resistance due to the constrictions at the both glottal and the constriction place as well as a variation in the supraglottal cavity volume induced by an expansion of the surrounding walls under the effect of the intra-oral pressure.

However, none of these studies include a self-oscillating model to represent the mechanical

behavior of the vocal folds which are modeled by a single parameter representing the average value of the glottal area [MW80]. As a consequence this model doesn't describe the cycle to cycle oscillation of the vocal folds nor can predict the onset and offset of voicing. Voicing is assumed to occur when the difference pressure across the glottis reaches an arbitrary critical value (typically of order of 200 Pa) [Kea84, WK86].

A more detailed approach has been proposed by McGowan et al. [MKL95] where a two-mass model for the vocal folds was added to produce complete aerodynamic simulation of vowel - consonant - vowel sequences. In the case of the production of an /apa/ sequence, for instance, McGowan et al. [MKL95] concluded that, in order to predict an oral volume flow velocity comparable to the one measured on a human subject, a large number of parameters need to be precisely controlled during the simulation. These parameters were respectively, the sub-glottal pressure, the glottal rest area, the stiffness of the vocal folds, the constriction area and an extra volume flow velocity due to the articulatory movement. Complex time contours of these parameters were also needed, with typical variations of order of 20 ms. However, as mentioned by Stevens [Ste93], some of these effects could not be articulatory induced but rather generated by a passive phenomenon due aerodynamics.

The present paper is in the vein of these previous studies aiming at a better understanding of the physics of vowel - plosive sound production. By refining existing theories, a particular interest will be paid on the aerodynamic interaction between the vocal folds, the constriction generated during the plosive and supraglottal walls, as this phenomenon is often reported to be of great importance [Kea84, WK86, Ste93].

The consequences of these phenomena on voicing onset and offset as well as passive changes of the glottal area will be evaluated. Compared with the above mentioned studies, our approach involves a validation against experiments on a mechanical replica prior to the comparison with speech data.

This work will focus on bilabial plosives since this production can be easily and extensively studied and measured *in vivo*. The approach is thus first, to collect articulatory, aerodynamic and acoustic data on a human participant in order to obtain a typical example of production. Then, theoretical models to explain the measured data are proposed. At first, a model for the airflow through the lips will be discussed, then a model for the deformation of the vocal tract walls will be presented. The relevance and the accuracy of these models are evaluated against measurements on a mechanical replica of the phonatory system. Lastly, numerical simulations are realized and compared with the measured data on the speaker.

6.2 Measurements on a human speaker

6.2.1 Example of an /apa/ sequence

The speaker was a French male adult participant who was asked to utter a /VCV/ sequence made of the French [a] vowel (V) and the French voiceless bilabial [p] consonant (C). Measurements were performed using acoustic and flow transducers synchronized with a high speed video camera. First, the acoustic pressure P_a is measured by the means of a cardioid microphone (AKG, C1000S), placed in front of the subject's lips at a distance of 1 m. The

intra-oral pressure IOP , is simultaneously measured using a probe tube of length 1.2 m and of diameter 5 mm, inserted inside the mouth through the lip corner and linked to a differential pressure sensor (Kulite XCS093). The pressure sensor was calibrated against a water column manometer with a typical accuracy of ± 5 Pa.

A high speed camera (Mikrotron, EoSens Cube7), equipped with a macro lens (Tokina, AT-X100mmf/2.8 PRO D), recorded the lip motion at a rate of 500 Hz. The video signal was synchronized with the pressure signals using a trigger sent from the A/D converter (National Instruments, PCI-MIO-16XE-10) to the camera. The lip parameters, illustrated in the figure 6.1, namely the lip width A and the lip aperture B , were obtained from the video recording.

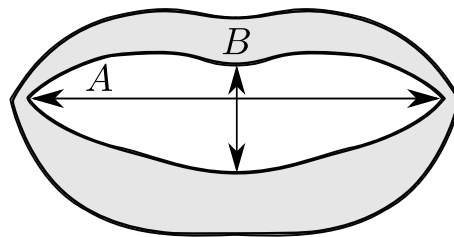


FIGURE 6.1 – Schematic of the lips : the lip width A and lip aperture B are the length of the main axis of the internal contour of the lips.

These parameters were extracted from each frame of the video recording, using a semi-manual method. The inner lip contour, computed using Canny edge detector [Can86], was first plotted over the picture. Then, the lip parameters were measured by detecting manually each corner of the lip area. Prior to the recordings, a picture of the participant holding a 4 mm long benchmark grid (Edmunds, Optics, 58-607) between the lips was realized in order to perform the conversion from pixels to centimeters. The accuracy on the estimation of both A and B was ± 0.15 mm.

The productions were verified *a posteriori* using the recording, i.e. the acoustic realization, the spectrogram and the intra-oral pressure measurement.

An example of measurement is presented in the figure 6.2. For the sake of clarity, the pressure fluctuations due to the acoustics are removed from the IOP signal, using a low-pass filter. As expected from previous studies [MW80, Wes83, Kea84, WK86, Dar87] during the end of the lip closure phase, since the air flow from the lungs progressively decreases, the vocal folds stop to oscillate while the IOP increases up to a quasi-constant value. It can be observed from the acoustic signal that the phonation is sustained during a short period after the complete closure of the lips (of order of 40 ms). During the lip opening phase, the acoustic burst responsible for the plosive sound is synchronous with IOP decrease down to the atmospheric pressure. Those events are followed by the voicing onset of the second vowel.

One can observe a strongly asymmetrical pattern of the IOP increasing during 40 ms and decreasing within 15 ms although the lips opening and closing speed was found to be of the same order of magnitude, about 2 m.s^{-1} for A and 0.4 m.s^{-1} for B .

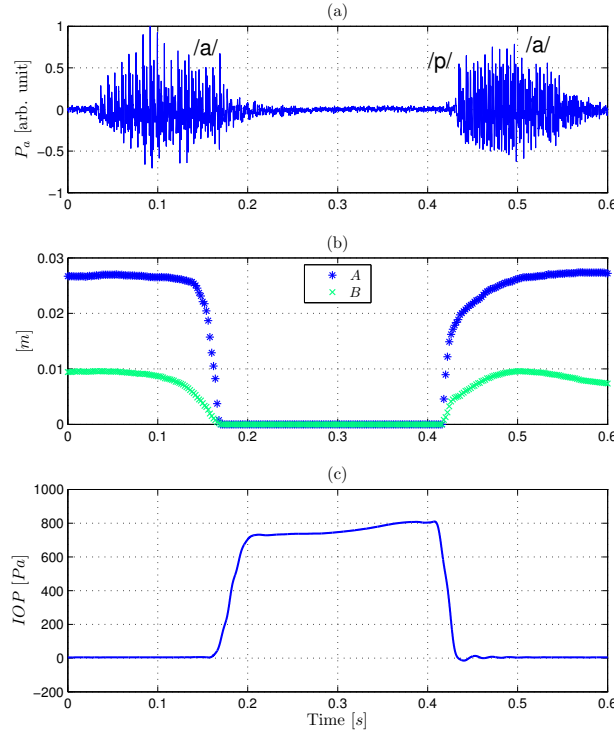


FIGURE 6.2 – Measurements for an /apa/ sequence of a human speaker, (a) : acoustic pressure P_a measured at 50 cm from lips, (b) : lip parameters (width A and height B) extracted from the video recording, (c) : intra-oral pressure IOP . [Listen to \$P_a\$ signal \(sound 11\)](#)

This example also enhances the need for a high speed video recording since about 75 % of the closing and opening movements are realized within 20 ms.

6.2.2 Influence of the cheeks expansion

In order to evidence and quantify the effects of an expansion of the supraglottal cavity, additional measurements were performed on the same speaker who was asked to place his hands on his cheeks to prevent them from inflating.

The intra-oral pressure was measured using the EVA2 work station [TG95]. The figure 6.3 shows an example of a comparison between the constrained (C) and unconstrained (U) conditions for the IOP during the closing phase. In the case where the oral cavity walls are constrained, the increase slope for IOP , averaged over 10 repetitions, is higher by 60 %.

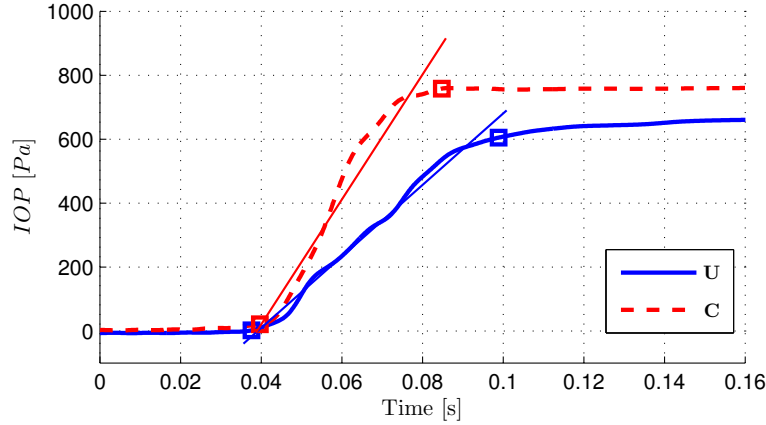


FIGURE 6.3 – Comparison of the increase in IOP during the production of an /apa/ sequence, under the constrained (C) and unconstrained (U) conditions.

6.3 Air flow model for the lip constriction

6.3.1 Theoretical models

Three theoretical models, used by Deverge et al. [DPV⁺03] to model the air flow across the glottis, are tested.

Bernoulli model

The first theoretical model is based on the Bernoulli equation for an ideal quasi-steady flow. As the area of the oral cavity is larger than the lip area, the kinetic energy upstream of the lips is neglected. The Bernoulli equation becomes :

$$p(x) = p(0) - \frac{\rho}{2} \left(\frac{\Phi_V}{W_\ell h(x)} \right)^2, \quad (6.1)$$

where $h(x)$ is the aperture along the coordinate parallel to the air flow, Φ_V is the volume flow velocity, ρ the constant air density and W_ℓ is the width of the lip channel. The pressure $p(x)$ is computed in two steps. First, Φ_V is determined by applying equation 6.1 between the beginning of the channel, for $x = 0$, and the abscissa of the flow separation point x_s . Then, the pressure $p(x)$ can be determined from Φ_V for any abscissa.

The aperture h_s at the flow separation point in the diverging part of the lips is predicted by the geometrical criterion used by Lous et al. [LHVH98], which links h_s to the minimum of the lip channel h_ℓ :

$$h_s = h(x_s) = \alpha h_\ell, \text{ with } \alpha = 1.1. \quad (6.2)$$

In the case of a steady flow, the accuracy of the prediction using this criterion is discussed by Van Hirtum et al. [HCP09].

Reynolds model

The second theoretical model uses the inviscid Bernoulli equation corrected by an additional term adapted from the lubrication theory of Reynolds and accounting for viscous losses [DPV⁺03]. The volume flow velocity Φ_V is determined by solving the quadratic equation :

$$\begin{aligned} p(x) &= p(0) - \frac{\rho}{2} \left(\frac{\Phi_V}{W_\ell h(x)} \right)^2 - \Delta P_\nu(x), \\ \text{with } \Delta P_\nu(x) &= \frac{12\rho\nu\Phi_V}{W_\ell} \int_0^x \frac{d\chi}{h^3(\chi)}, \end{aligned} \quad (6.3)$$

where ν is the kinematic viscosity coefficient.

Thwaites model When the viscous and the inertial terms are of the same order of magnitude, the boundary layer theory assumes that the flow can be divided into two regions. Within the first one, far enough from the walls, the flow can be considered as frictionless and one dimensional. In the second one, the boundary layer, the flow is considered as two-dimensional and viscous. Under the assumption of a thin boundary layer, the flow may be predicted using the Von Kàrmàn integral formulation [SG99]. The Von Kàrmàn equation is solved numerically using the Thwaites method [Vil02] which allows to predict the pressure distribution $p(x)$, the position of the flow separation point x_s and the thickness of the boundary layer.

6.3.2 Experimental validation on a replica of the phonatory system

Setup

In order to test the above mentioned theoretical flow models, the set-up illustrated in figure 6.4 was built. It consists of the following elements :

- a 0.6 m^3 reservoir (2) filled with acoustical foam, supplied by an air compressor (Atlas Copco, GA5 FF-300-8), acts as a constant pressure supply, controlled by a pressure regulator (Norgren, 11-818-987) allowing for input pressure in the range of $0 - 4000 \text{ Pa}$ (1),
- a self-oscillating replica of the vocal folds (3) is connected to the pressure reservoir using a rigid uniform tube of length 12 cm with an internal diameter of 2.5 cm ,
- a uniform tube of total length 18 cm including a 12 cm length plexiglass portion (4), having the same diameter, is placed downstream of the vocal folds replica to mimic

the vocal tract. For the measurements involving deformable walls the plexiglass tube is replaced by a latex tube,

- lastly, a replica of the lips (6) is connected to this later tube. The motion of the upper mechanical lip is controlled by a step motor.

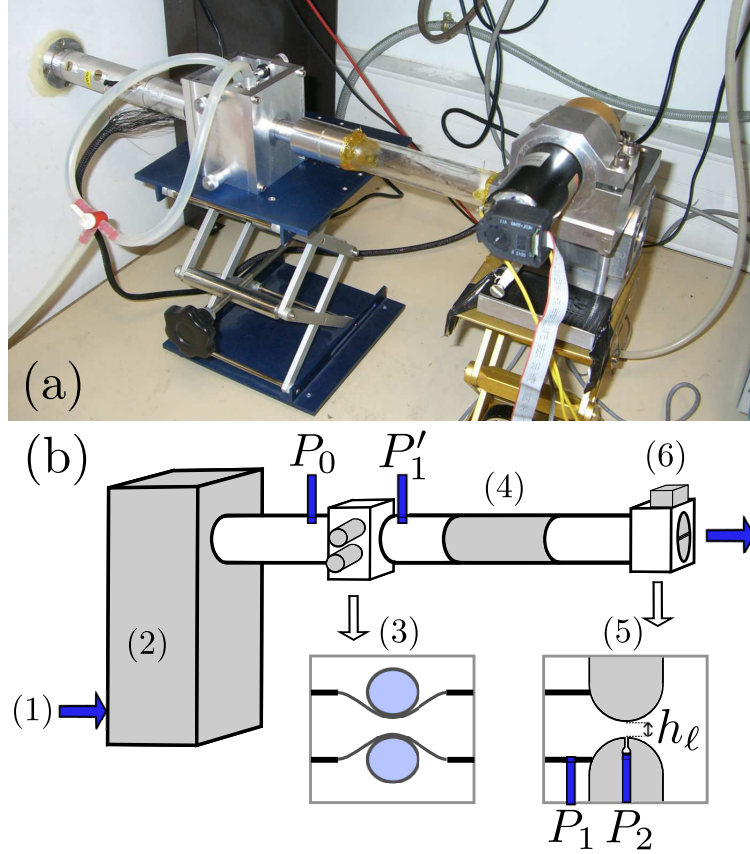


FIGURE 6.4 – (a) : Picture of the replica of phonatory system. (b) : Diagram of the replica of phonatory system including, (1) : air compressor and pressure regulator, (2) : pressure reservoir, (3) : replica of vocal folds, (4) : rigid or plexiglass tube, (5) : metal replica of lips, (6) : motor and mechanical lip displacement sensor.

The replica of the vocal folds is made of two latex tubes, filled with water under pressure. When placed transverse to an air flow, this replica is able to self-oscillate. More details on this vocal folds replica can be found in Pelorson and Laval [PL12]. The replica of the lips is made of two half metallic cylinders (Radius : $R = 1\text{ cm}$, width : $W_\ell = 3\text{ cm}$), placed perpendicularly to the flow direction. The geometry of the mechanical lip constriction is presented in figure 6.5. The minimum of lip aperture, for $x = R$, is noted h_ℓ . The height $h(x)$ of the lip constriction along the flow direction is thus given by :

$$h(x) = h_\ell + 2 \left(R - \sqrt{R^2 - (x - R)^2} \right). \quad (6.4)$$

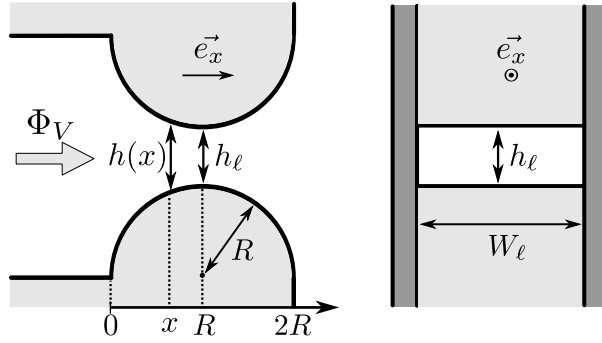


FIGURE 6.5 – Geometrical model for the lip constriction, side (left) and front (right) view.

To mimic the lip motion, a step motor imposes a sinusoidal perpendicular motion of the upper half cylinder (figure 6.5) with a period that could be controlled from 0.5 s up to 2.5 s. Because the lip replica surface roughness would prevent from complete closure, a latex layer of thickness 0.2 mm was glued to the mobile part of the mechanical lips. The measurement of the aperture h_ℓ was obtained using an optical sensor (Optek, OPB700). The non-linear behavior of this sensor required a calibration using pictures of the mechanical lip opening against a reference benchmark grid (Edmunds Optics, 58-607). The accuracy on h_ℓ is estimated to be ± 0.05 mm.

Four differential pressure transducers (Kulite, XCS-093-035-Bar-G) were placed along the setup to measure P_0 , P'_1 , P_1 at $x = 0$ and P_2 at $x = R$, which correspond respectively to the subglottal, supraglottal, intra-oral and intra-labial pressure according to figure 6.4. Prior to the experiments, the calibration of the pressure transducers was realized against a water-meter with an accuracy of ± 5 Pa. Pressure and optical signals were synchronized and recorded using an A/D converter (National Instruments, PCI-MIO-16XE-10), at a sampling frequency of 10 kHz.

Experiments are performed under two conditions : without and with vocal folds oscillations. For the measurement with vocal folds oscillations, the pressure of water inside the latex tubes is fixed to $8 \cdot 10^3$ Pa which provides stable oscillation at a frequency of 104 ± 2 Hz. For the measurement without vocal folds oscillations, the water pressure is lowered to $4.5 \cdot 10^3$ Pa as to deflate the vocal folds replicas.

Results

The relevance of the theoretical models to represent the air flow between the lips is evaluated by comparing their predictions with the experimental data.

At first, the joint evolution of pressure P'_1 and P_1 was systematically investigated. It was found that the maximum difference between these two quantities, reached when the lips were completely closed, was less than 2 %. This result was to be expected since viscous losses scale with the inverse third power of the cross section area and thus become negligible within the vocal tract replica.

Figures 6.6 present one example of the obtained results in the case of non oscillating vocal folds while figures 6.7 present the same experiment realized with oscillating vocal folds.

Both cases, with and without vocal folds oscillations, lead to the same conclusions. The Bernoulli model appears to be useless since it always predicts a negative pressure for P_2 for any aperture h_ℓ which is not observed experimentally.

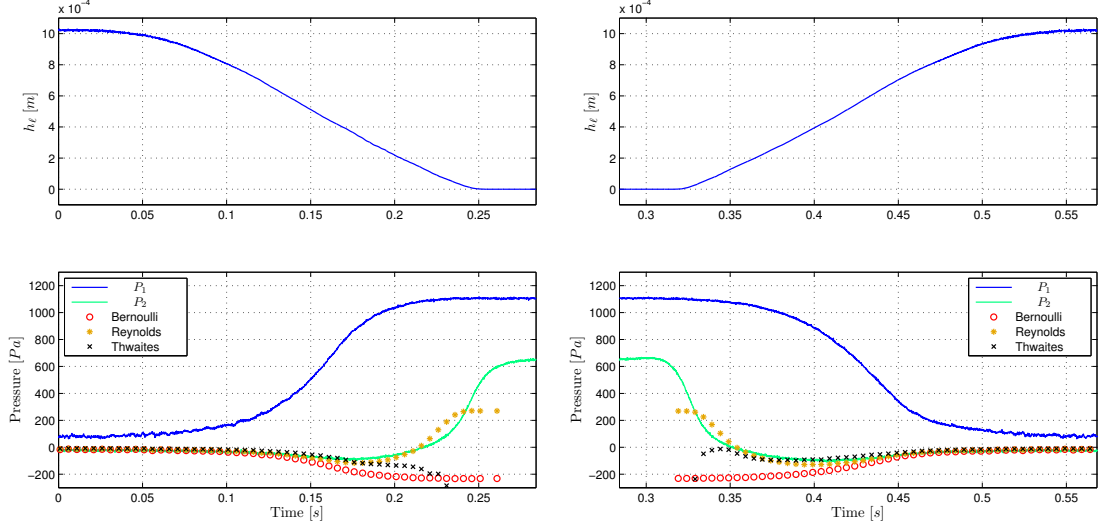


FIGURE 6.6 – Experimental and theoretical results for $x = R$, in the case of non oscillating vocal folds, during the mechanical lip closure (left) and opening (right), for a sine wave lip motion with a period of 0.5 s.

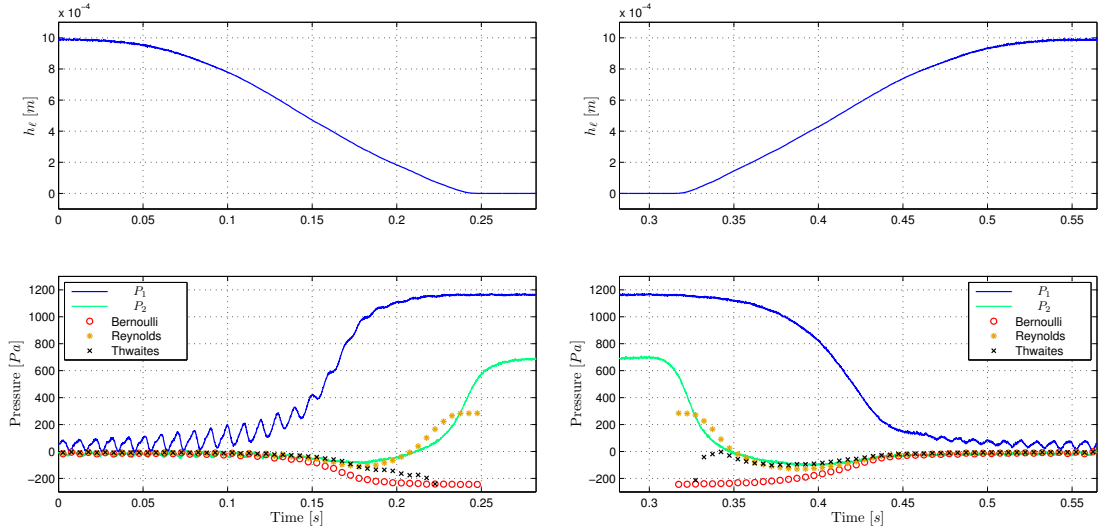


FIGURE 6.7 – Experimental and theoretical results for $x = R$, in the case of oscillating vocal folds, during the mechanical lip closure (left) and opening (right), for a sine wave lip motion with a period of 0.5 s.

The prediction of the Thwaites method shows a good agreement with the measured pressure for $h_\ell > 0.1$ mm during the opening phase but only for $h_\ell > 0.25$ mm in case of the closing phase. This difference is likely to be due to an unsteady effect caused by the

motion of the lip channel walls. The Thwaites method which is based on the boundary layer theory fails thus to explain the experimental data for the smallest mechanical lip apertures which are those of interest for plosive sound production. This departure is due to the fact that, when h_ℓ tends to 0, viscous effects become predominant and cannot be considered as small on average. As the matter of fact, the Reynolds flow model seems to be the most adapted theoretical model to describe the air flow in the lips for both conditions.

6.4 Aerodynamic model of the supraglottal cavity expansion

In section 6.2.2, it was shown that the expansion of the supraglottal cavity, induced by the pressure upstream to the constriction, was leading to a significant variation of the *IOP*. In this section, a theoretical model to account for this effect is introduced. The model is evaluated against measurement made on the replica of the phonatory system.

6.4.1 Theoretical models

The theoretical models presented here include a self-oscillating vocal folds model coupled to a labial constriction. The aerodynamic and mechanical models of the vocal folds and the air flow model for the lip constriction are exposed first, then the model proposed to describe an expansion in the supraglottal cavity is presented. In the following the subscripts "g", "w" and "ℓ" will refer to the glottis, the walls of the vocal tract and the lips, respectively.

Air flow and mechanical models for the vocal folds

The two mass model described by Lous et al. [LHVH98], which will be briefly described in this section, is used as the theoretical model for the vocal folds oscillation. The geometry of the glottal constriction is represented by three plates located between abscissas x_i and x_{i+1} with $i = 0, 1, 2$, as shown in figure 6.8.

The simulation is driven by the subglottal pressure P_{sub} . Based on previous observations made during measurement on the replica (see section 6.3.2), the intra-oral pressure P_{io} is assumed to be equal to the supraglottal pressure. In this study, the effects of acoustical coupling on pressure difference across the glottis, upstream and downstream of the larynx, are not taken into account.

Within the glottis, the pressure along the glottis is given by the following equation :

$$p(x, t) = \begin{cases} P_{sub}(t) - \frac{\rho}{2} \left(\frac{U_g(t)}{W_g} \right)^2 \left(\frac{1}{h^2(x, t)} - \frac{1}{h_0^2(t)} \right) - \frac{12\rho\nu U_g(t)}{W_g} \int_{x_0}^x \frac{d\chi}{h^3(\chi, t)}, & \text{for } x_0 \leq x < x_{gs}, \\ P_{io}(t), & \text{for } x \geq x_{gs}, \end{cases} \quad (6.5)$$

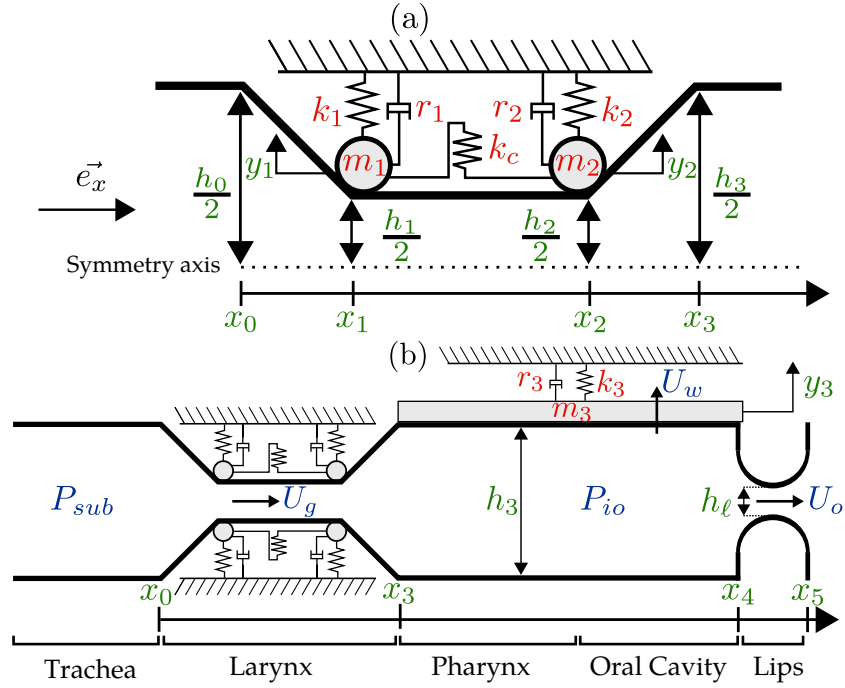


FIGURE 6.8 – (a) : diagram representing the two-mass model of vocal folds, including the geometrical and the mechanical parameters, (b) : diagram representing the model of the phonatory system with a deformable vocal tract, including the aerodynamic and geometrical parameters.

where U_g is the volume flow velocity, W_g is the width of the glottal channel and x_{gs} is the abscissa at which the flow separates from the glottis to form a jet. The position of the flow separation point is obtained using a geometrical criterion similar to the one presented in section 6.3.1 :

$$h_{gs} = h(x_{gs}) = \min(\alpha h_1, h_2), \quad (6.6)$$

with $\alpha = 1.1$.

The equations of motions of each mass m_i is :

$$m_i \ddot{y}_i + r_i \dot{y}_i + k_i y_i + k_c (y_i - y_j) = F_i, \quad (6.7)$$

for $i, j = 1, 2$ and $i \neq j$.

The parameters k_i and r_i are respectively the stiffness and the damping coefficient of the spring attached to mass m_i . k_c is the coupling stiffness between mass m_1 and mass m_2 , y_i are the displacements of each mass from their rest positions y_0 relative to the symmetry axis.

The instantaneous glottal aperture $h_i(t)$ at abscissa x_i is linked to the displacement y_i by the relationship $h_i(t) = 2(y_i(t) + y_0)$, for $i = 1, 2$. The aerodynamic forces $F_i(t)$ applied to each mass m_i of the vocal folds model are given by :

$$F_i(t) = W_g \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} p(x, t) dx + W_g \int_{x_i}^{x_{i+1}} \frac{x_{i+1} - x}{x_{i+1} - x_i} p(x, t) dx, \quad (6.8)$$

for $i = 1, 2$.

When the vocal folds collide, i.e. when $h_i(t) \leq 0$, the stiffness k_i is arbitrary multiplied by 4 while the damping coefficient r_i is set to the critical value $2\sqrt{m_i k_i}$. Both U_g and P_{io} become 0, the only remaining force being the static one due to the pressure P_{sub} upstream of the contact point of the vocal folds. The equations of motion are solved numerically for displacements y_i using the backward Euler method with a sampling rate of $3 \cdot 10^4$ Hz.

Air flow model within the lips

The geometrical model used to represent the lip constriction is the one presented in figure 6.5. Following the conclusions of section 6.3, the Reynolds model is adopted for the description of the flow. Applied up to the flow separation point, for $x = x_{\ell_s}$, this yields :

$$P_{io}(t) = \frac{\rho}{2} \left(\frac{U_o}{W_\ell h_{\ell_s}} \right)^2 + \frac{12\rho\nu U_o}{W_\ell} \int_{x_4}^{x_{\ell_s}} \frac{d\chi}{h^3(\chi)}, \quad (6.9)$$

where U_o is the volume flow velocity coming out from the oral cavity.

Aerodynamic and mechanical model for the supraglottal cavity expansion

Following Rothenberg's model [Rot68], the deformation of vocal tract walls is represented by a spring-mass-damper system, as shown in figure 6.8. The motion of this deformable wall can therefore induce an additional positive or negative volume flow velocity U_w . The interaction between this phenomenon and other aerodynamic quantities can be written as a system of three equations :

$$\begin{cases} A_w P_{io}(t) &= m_3 \ddot{y}_3(t) + r_3 \dot{y}_3(t) + k_3 y_3(t), \\ U_w(t) &= A_w \dot{y}_3(t), \\ U_g(t) &= U_w(t) + U_o(t), \end{cases} \quad (6.10)$$

where k_3 , r_3 and y_3 are respectively the equivalent stiffness, damping coefficient and displacement with respect to the rest position of the mass m_3 and A_w is the area of deformable wall of supraglottal cavity.

Together with the equations of the flow across the glottis (equation 6.5) and across the lips

(equation 6.9) one obtains thus a system of 5 equations with 5 unknowns, which are P_{io} , U_g , U_o , U_w and y_3 . In the case of a closed glottis, equation 6.5 is removed from the system and U_g is set to zero. In the same way, the equation 6.9 is removed and $U_o = 0$, when the lips are closed. All time derivative terms are computed numerically using backward Euler method.

6.4.2 Experimental validation by numerical simulations

Measurements on a replica of the phonatory system

The validation of the theoretical model was made against measurements on the replica of the phonatory system presented in section 6.3.2. The plexiglass tube representing the vocal tract was replaced with a deformable tube made of latex, having a thickness $e = 0.2 \text{ mm}$, a length $L_t = 7 \text{ cm}$ and an internal diameter of 3 cm , to mimic the passive expansion of human tissues within the supraglottal cavity.

The mechanical parameters of the mass-spring-damper system equivalent to the deformable tube were experimentally estimated as follows. The mass of the latex tube m_3 was simply measured using a weighing scale with an accuracy of $\pm 0.1 \text{ g}$.

For the determination of k_3 , the high speed camera, at a rate of 500 Hz , was used to measure the volume variation ΔV_t of the latex tube during the closure and the opening of the mechanical lips. The external diameter d_{ext} along the tube length was extracted from the video recordings using an automatic image processing. The conversion from pixels to meters was realized using a picture of a ruler placed over the latex tube, in the same focal plane with an accuracy $\pm 0.1 \text{ mm}$. Assuming that the shape of the latex tube is a barrel whose revolution axis is parallel to the flow direction, the volume V_t inside the latex tube is given by :

$$V_t(t) = \int_0^{L_t} A_{int}(x, t) \, dx = \frac{\pi}{4} \int_0^{L_t} (d_{ext}(x, t) - 2e)^2 \, dx. \quad (6.11)$$

This volume V_t could therefore be computed from d_{ext} , for each video frame, using a numerical integration.

Figure 6.9 presents the measured relationship between P_1 and the volume variation ΔV_t of the latex tube, defined as the difference between $V_t(t)$ and the initial volume of the tube (at rest), during an opening and closing of the mechanical lips. Both quantities can be easily linked to the force $P_1 A_w$ applied to the mobile mass m_3 and its displacement y_3 such as $y_3 = \Delta V_t / A_w$. Although a slight hysteresis can be observed, the relationship between P_1 and ΔV_t is quasi-linear. Assuming that the inertial force is negligible with respect to the elastic one, the equivalent stiffness k_3 of the system can be estimated using a linear regression which yields : $k_3 = \beta A_w^2$, where β is the obtained slope coefficient defined by $\Delta V_t = \beta P_1$. The area A_w is estimated considering the inner surface area of the latex tube. Lastly, the damping coefficient is chosen in order to have a critically damped system. The table 6.1 summarizes the parameters of the equivalent mass-spring-damper system.

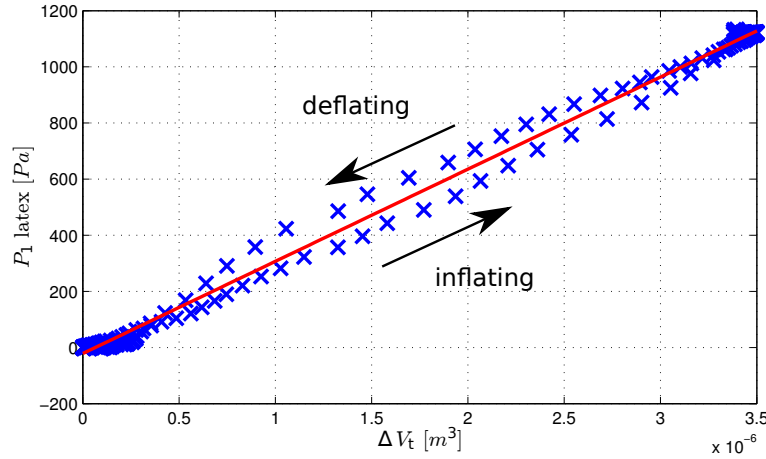


FIGURE 6.9 – Estimation of the equivalent spring constant : pressure P_1 upstream to the lips as a function of the volume variation ΔV_t of the latex tube. The crosses represent the experimental data and the line represents the linear regression made to estimate the equivalent stiffness k_3 .

Stiffness	$\beta = 3.28 \cdot 10^8 \text{ Pa.m}^{-3}$ $A_w = 6.6 \cdot 10^{-3} \text{ m}^2$ $k_3 = \beta A_w^2 = 1.43 \cdot 10^4 \text{ N.m}^{-1}$
Mass	$m_3 = 1.6 \cdot 10^{-3} \text{ kg}$
Damping coefficient	$r_3 = 2\sqrt{m_3 k_3} = 9.57 \text{ kg.s}^{-1}$

TABLE 6.1 – Parameters of mass-spring-damper system to model the expansion of the latex tube.

Results

The input parameters of the simulation are the measured mechanical lip aperture h_ℓ and the supply pressure P_0 . Figure 6.10 presents the pressure P_1 measured together with the result of the fitted simulation in case of a rigid and a deformable tube. As expected, using this vocal tract wall expansion model, the increase in simulated pressure P_1 during lip closure is less steep.

The maximum level of the simulated volume variation is in agreement with the measured one. The difference between these two quantities is below 1%. However, simulation and measurements present some differences in the evolution of ΔV_t and P_1 , especially during and just after the lip closure phase.

This departure between measurement and simulation may be due to the aerodynamic prediction. Results of measurements presented in section 6.3.2 showed that the Reynolds flow model underestimates the pressure in the mechanical lip constriction. Turbulent flow or unsteady effects due to the flow induced by the motion of the lips could be responsible for this departure.

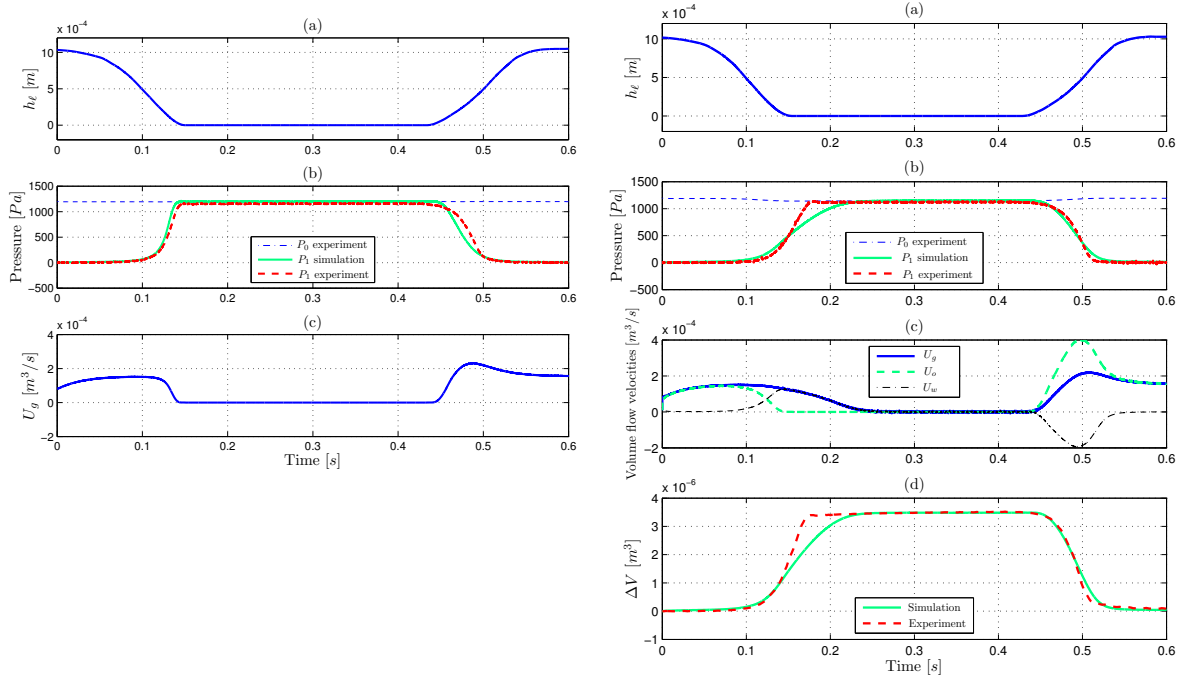


FIGURE 6.10 – Numerical simulations of the replica measurement in the rigid (left) and in the deformable vocal tract configuration (right), (a) : measured aperture h_ℓ of the mechanical lips : input parameter of the simulation, (b) : measured pressure P_0 : input parameter of the simulation, measured and simulated pressure P_1 , (c) : simulated glottal volume flow velocity U_g , (d) : comparison between measured and simulated supraglottal cavity volume ΔV .

6.5 Numerical simulations of an /apa/ sequence

In this section, the measurements obtained during an /apa/ sequence (figure 6.2) is reproduced using the numerical simulation based on the theoretical models presented in section 6.4.1.

6.5.1 Parameters of the simulations

The subglottal pressure P_{sub} is fixed constant and equals the maximum value of the measured *IOP*. The vocal folds parameters proposed by Lous et al. [LHVH98] have been chosen and are kept constant, no laryngeal adjustments are assumed. The length $x_5 - x_4$ of the lip constriction is constant which means that any protrusion movement is neglected. The only time-varying parameters during the simulation are thus the measured lip geometrical parameters A and B . In order to obtain signals at the frequency sampling of the simulation (30 kHz), these parameters are interpolated from those measured on the human speaker. The parameters of lip channel model are chosen in order to keep the same aspect ratio and the same labial area, approximating the internal contour of the human lips by an ellipse [PHRB97], it yields $W_\ell = A\sqrt{\pi/4}$ and $h_\ell = B\sqrt{\pi/4}$.

The estimation of parameters of the mass-spring-damper system equivalent of the deformable vocal tract walls is performed in the following way. The stiffness k_3 is determined by fitting the *IOP* curve, assuming that $A_w = 0.01 \text{ m}^2$. The mass m_3 is approximated by an estimation of the vibrating volume in supraglottal cavity walls, considering that the density of human tissues is equivalent to the density of water and that the thickness of the tissues involved in the expansion is $e_w = 5 \text{ mm}$. The damping coefficient r_3 is set to its critical value. All of these values are summarized in table 6.2. Note that these parameters are of the same order of magnitude than those used to simulate the measurements on the mechanical replica.

Geometrical parameters	Mechanical parameters
$h_0 = 1.79 \text{ cm}, h_3 = 1.86 \text{ cm}$	$m_1 = m_2 = 0.1 \text{ g}$
$y_0 = 8 \cdot 10^{-5} \text{ m}, W_g = 1.4 \text{ cm}$	$k_1 = k_2 = 60 \text{ N.m}^{-1}, k_c = 0.6 \text{ k}$
$x_1 - x_0 = 0.2 \text{ mm}$	$r_1 = r_2 = 0.1 \cdot 2\sqrt{m_1 k_1}$
$x_2 - x_1 = 2 \text{ mm}$	$m_3 = \rho_{wat} A_w e_w = 5.0 \cdot 10^{-2} \text{ kg}$
$x_3 - x_2 = 0.2 \text{ mm}$	$(\rho_{wat} = 1000 \text{ kg.m}^{-3})$
$x_5 - x_4 = 1 \text{ cm}$	$k_3 = 1.3 \cdot 10^4 \text{ N.m}^{-1}$
$A_w = 0.01 \text{ m}^2, e_w = 5 \text{ mm}$	$r_3 = 2\sqrt{m_3 k_3} = 13.27 \text{ kg.s}^{-1}$

TABLE 6.2 – Geometrical and mechanical parameters of the two-mass model, used for the simulation of an /apa/ sequence.

6.5.2 Results

In this study, the intra-oral pressure appears to be a relevant quantity to compare the simulation with the measurements because, since the self-oscillation of vocal folds is driven by the pressure difference across the glottis $\Delta P = P_{sub} - P_{io}$, predicting the intra-oral evolution allows to predict the voice onset and offset times during the production of a plosive. The simulations presented here attempt to reproduce the example showed in figure 6.2.

Figure 6.11 presents results of simulations in the case of a rigid and deformable vocal tract.

Evolution of P_{io}

For a rigid vocal tract simulation, compared with speech data, the slope of increase in intra-oral pressure is significantly higher by about 580 %. A similar effect but less important is observed during the lip opening phase : the simulation predicts a steeper decrease of the intra-oral pressure by 360 %.

In the case of a deformable vocal tract simulation, the intra-oral pressure P_{io} presents a much better agreement with the measured one in particular during the lip closure. The slopes agree within 30 %. During the lip opening phase, the vocal tract walls movement lowers the slope of the intra-oral pressure drop. However, the pressure drop remains steeper by 90 % than the one measured on the reference speaker.

Effect of supraglottal cavity expansion on phonation

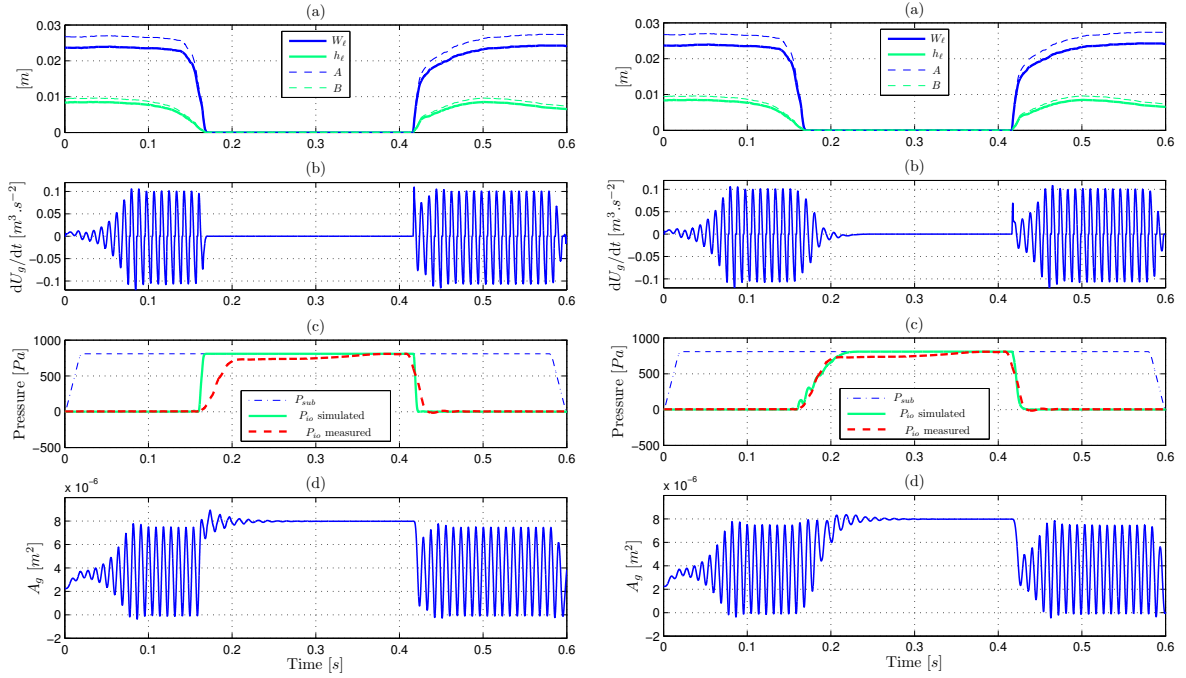


FIGURE 6.11 – Numerical simulations for a vowel-bilabial plosive-vowel sequence using rigid vocal tract model (left) and the vocal tract expansion model (right), (a) : polynomial interpolation W_ℓ and h_ℓ of lip parameters A and B measured on a human speaker and represented by the markers, (b) : time derivative dU_g/dt of the volume flow velocity simulated, (c) : subglottal pressure P_{sub} , simulated intra-oral pressure P_{io} and measured intra-oral pressure IOP , (d) : simulated glottal cross-sectional area A_g defined as $A_g = W_g \min(h_1, h_2)$.

During the closed phase of the lips, the glottal volume flow velocity U_g becomes zero because of the equalization of the intra-oral and the subglottal pressure. Thus, the only remaining force is due to the static pressure P_{sub} which is exerted on the vocal folds and thus affects the glottal area A_g . The simulated glottal area A_g was thus found to increase from an initial value of $A_g = 2.24 \cdot 10^{-6} \text{ m}^2$ at $t = 0$ to a larger one, $A_g = 8.06 \cdot 10^{-6} \text{ m}^2$ for $0.25 < t < 0.42 \text{ s}$.

In the case of a rigid vocal tract, the simulation of dU_g/dt exhibits an abrupt voice offset starting at the time when the lips are completely closed. The simulations considering a deformable vocal tract lead to a less abrupt behavior, the vocal folds having an oscillatory motion for several periods. Thus, the phonation is sustained during 40 ms after lip closure.

6.6 Conclusion

Aerodynamic, acoustic and video recording of the lips were made on a human speaker during the production of an /apa/ sequence in order to quantify accurately the different time patterns of the measured signals. On this basis, theoretical models to describe the air flows between the lips and the effect of an expansion of the supraglottal cavity are proposed and

validated against measurement on a replica of phonatory system, including a self-oscillating replica of the vocal folds, a vocal tract and a controlled replica of the lips. Finally, numerical simulations of an /apa/ sequence were realized and compared to measurement on a human speaker, using intra-oral pressure.

The conclusions that can be drawn from this study are the followings.

- A steady and viscous model for the flow through the lips seems to be a reasonable approximation especially during the opening and the closure of the lips which are the time intervals of interest for plosives.
- Simulations based on a two-mass model suggest that the realization of an occlusion can interact significantly with the vocal folds. This study confirms the conclusion of Stevens [Ste93] since the voice offset and the increase in the glottal area are predicted during or just after lip closure, without the need to adjust the glottal control parameters.
- Modeling the passive expansion of the supraglottal cavity allows to obtain intra-oral pressure in better agreement with the one measured on a subject. This phenomenon also explains the extension of phonation up to 40 *ms* after the closure of the lips, whereas, after the lips opening, the effect on the voice onset time is not significant.
- Although simulations are realized using only the constriction area as a time varying parameter, the results show reasonably agreement with the measured data on a speaker. Some residual departures might however be explained by the need for a more precise time control of the control parameters, as done by McGowan et al. [MKL95].

Our future work would rather focus on attempting to improve the physical models. Acoustic coupling with the vocal tract and the trachea was, for instance, neglected in this study. This phenomenon can however have an influence on the phonation threshold pressure [LLH⁺12] and thus, potentially, on voice onset and offset.

Finally, although this study concentrated on bilabial plosives for practical reasons, the mechanical and aerodynamic models developed in this work should be relevant and will be extended to plosives made with other places of occlusion within the vocal tract.

Influence du couplage acoustique sur le *Voice Onset Time*

Sommaire

7.1 Introduction	133
7.1.1 Éléments bibliographiques	134
7.1.2 Objectif de l'étude	134
7.2 Mesures sur maquette de l'appareil phonatoire	136
7.2.1 Dispositif expérimental	136
7.2.2 Méthode d'analyse des signaux mesurés	136
7.2.3 Résultats	138
7.3 Simulations numériques	138
7.3.1 Réalisation des simulations numériques	139
7.3.2 Méthode d'analyse des simulations	140
7.3.3 Résultats	140
7.4 Discussion et conclusion	141
7.4.1 Discussion	141
7.4.2 Conclusion	143

7.1 Introduction

Les travaux réalisés dans ce chapitre se basent sur la modélisation aérodynamique de l'écoulement dans les lèvres, présentée au chapitre 6. Dans cette étude, le couplage acoustique est pris en compte et utilisé afin d'étudier son influence sur le VOT (cf. définition en section 1.3.4) lors de la production d'une séquence voyelle - plosive non voisée - voyelle. La première partie de ce chapitre (section 7.1) est consacrée aux éléments bibliographiques nécessaires à la formulation de l'objectif et des hypothèses de l'étude. Les sections 7.2 et 7.3 présentent respectivement les mesures réalisées sur la maquette de l'appareil phonatoire ainsi que les simulations.

7.1.1 Éléments bibliographiques

7.1.1.1 Notion de seuil dynamique

À l'aide d'un modèle simplifié de clarinette, Bergeot et al. [BAVG13] ont mis en évidence l'existence d'un seuil de pression dynamique théorique, qui caractérise la valeur de la pression dans la cavité orale du musicien à l'apparition des oscillations dans le cas où la pression varie dans le temps. Ainsi pour une augmentation de la pression d'alimentation, la pression de seuil dynamique est supérieure à la pression de seuil statique. Ce phénomène est interprété comme la manifestation d'un retard de bifurcation (de Hopf).

Ces résultats impliquent que les seuils de pression déterminés expérimentalement en augmentant la pression sous-glottique sont sur-estimés par rapport aux seuils de pression calculés à partir de théories statiques avec des méthodes basées sur la linéarisation du système [CGC00, Rut07, LLH⁺12, Her14].

Bergeot et al. [BAVG13] montrent également que, dans le cas de simulations numériques, la sur-estimation de la pression de seuil est en partie limitée lorsque par de faible précision de calcul numérique.

7.1.1.2 Influence du contexte phonétique sur le VOT

De nombreuses études phonétiques ont montré qu'en moyenne le VOT (défini en section 1.3.4) augmente lorsque le lieu d'articulation de l'occlusion est déplacé des lèvres vers le voile du palais et ce, quelque soit le contexte vocalique ou le langage utilisé [Kla75, Zue76, CH88, Byr93, Yao09]. Les études phonétiques s'accordent également à dire que le VOT est généralement plus long lorsque la plosive est suivie par une voyelle fermée (par opposition à une voyelle ouverte) [FJ72, Kla75].

7.1.2 Objectif de l'étude

Bien que de nombreuses études portent sur l'influence du contexte phonétique ou bien des caractéristiques phonétiques liées au langage du locuteur sur le VOT, et que l'influence des phénomènes physiques sur la pression de seuil des oscillations soit relativement bien caractérisée (section 1.2.3), peu de travaux cherchent à établir un lien entre ces deux axes de recherche.

En partant du principe que les oscillations apparaissent au-delà d'un seuil de pression trans-glottique ΔP (différence de pression de part et d'autre de la glotte), Koenig et al. [KFL11] étudient l'effet du contexte phonétique sur la valeur de la pression P_{io} à l'extinction et à l'apparition des oscillations, dans le cas de la production de consonnes non voisées. Cette étude montre notamment que pour les consonnes suivies d'une voyelle fermée (/i/), la valeur de P_{io} à la reprise des oscillations est plus importante que celle mesurée dans le cas d'une voyelle ouverte (/a/). L'interprétation proposée est que la pression trans-glottique de seuil ΔP_{on} est plus faible dans le cas de la voyelle fermée en raison de l'interaction acoustique plus

importante avec le conduit vocal. L'effet d'un éventuel retard à la bifurcation n'est cependant pas considéré dans cette étude qui s'intéresse à l'aspect dynamique de la production. Les travaux présentés dans ce chapitre se basent sur les hypothèses suivantes : pendant la production d'une séquence voyelle - plosive non voisée - voyelle, la pression P_{sub} est constante (ce que ne considèrent pas Koenig et al. [KFL11]), et aucun d'ajustement laryngé n'est réalisé (cf. chapitre 6). Ces hypothèses impliquent que l'apparition des oscillations après le relâchement de l'occlusion est uniquement régie par l'évolution de la pression trans-glottique ΔP . L'idée proposée dans ce chapitre est que le retard à la bifurcation (cf. section 7.1.1.1) s'applique également au cas des plosives. Ainsi après le relâchement de l'occlusion, la pression trans-glottique augmente rapidement, le seuil statique ΔP_{St} est atteint puis les oscillations apparaissent avec un retard Δt . La figure 7.1 illustre la transposition de la notion de retard à la bifurcation au cas des plosives.

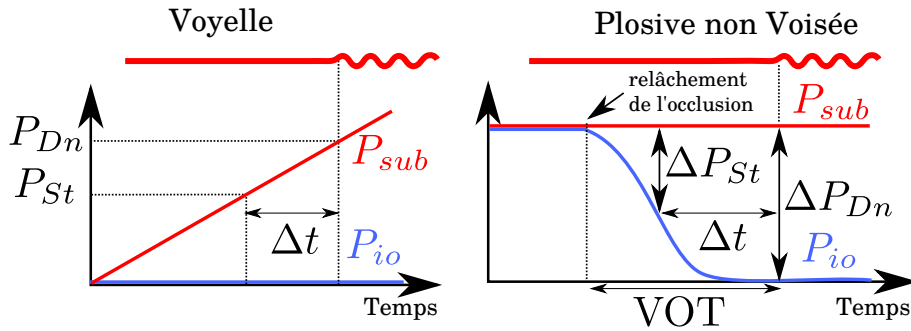


FIGURE 7.1 – Illustration de l'apparition retardée des oscillations dans le cas de la production d'une voyelle et d'une plosive non voisée. Les indices St et Dn désignent respectivement les seuils statique et dynamique.

Ce chapitre s'intéresse au lien entre le VOT et les notions de pression de seuil dynamique et de retard à la bifurcation. La littérature mentionne que, dans le cas où le couplage acoustique reste faible, la pression de seuil P_{on} diminue lorsque F_1 diminue, se rapprochant ainsi de F_0 (section 1.2.3). En supposant qu'une faible valeur de P_{on} entraîne, après le relâchement de l'occlusion, une apparition des oscillations plus rapide, une diminution de F_1 pourrait être responsable d'une valeur de VOT plus faible. Ce phénomène pourrait expliquer les tendances observées dans les études phonétiques (section 7.1.1.2).

L'objectif de ce chapitre est vérifier cette hypothèse par le biais de mesures réalisées sur la maquette de l'appareil phonatoire, puis en effectuant des simulations numériques. Les simulations permettront d'évaluer la capacité des modèles théoriques utilisés à reproduire les phénomènes observés.

7.2 Mesures sur maquette de l'appareil phonatoire

7.2.1 Dispositif expérimental

Le dispositif expérimental utilisé pour étudier l'influence du couplage acoustique sur le VOT est constitué de la maquette de l'appareil phonatoire présentée en section 4.2.1 et de la réplique de lèvres mécaniques (section 6.3.2). Pour cette étude, le résonateur amont est un tube en métal d'une longueur de 21,5 cm. La longueur du résonateur aval est modifiée en utilisant des tubes en plexiglas de différentes longueurs L_p variant de 4 à 22 cm (cf. figure 4.1) et un adaptateur métallique de 12,5 cm de long. Cette configuration est présentée en figure 7.2.

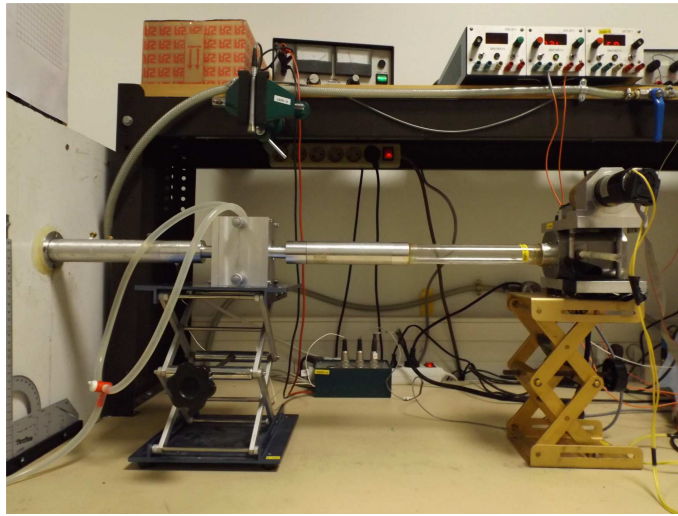


FIGURE 7.2 – Photographie de la maquette de l'appareil phonatoire avec, en aval de la maquette de cordes vocales, un résonateur composé de l'adaptateur en duralium d'un tube de plexiglass de 16 cm de longueur.

La pression statique de l'eau dans les répliques de cordes vocales est fixée à 5500 Pa. Les signaux mesurés sont la pression P_0 en amont de la réplique de cordes vocales et la pression P_1 en amont de la maquette de lèvres. Les capteurs, le matériel et les réglages sont identiques à ceux présentés en section 4.2.

7.2.2 Méthode d'analyse des signaux mesurés

Pour chaque valeur de L_p , le VOT est déterminé à partir des mesures des instants t_{ouv} et t_{on} qui correspondent respectivement aux instants de l'ouverture des lèvres mécaniques et de l'apparition des oscillations. La figure 7.3 montre un exemple de signaux mesurés pour $L_p = 10$ cm. L'oscillation des répliques de cordes vocales est interrompue par la fermeture des lèvres mécaniques en raison de l'augmentation de P_1 jusqu'à la valeur de P_0 . Après le relâchement de l'occlusion qui se produit pour $t_{ouv} = 500$ ms, les oscillations auto-entretenues

réapparaissent progressivement à partir de $t_{on} = 850 \text{ ms}$. Il est important de noter que les oscillations apparaissent alors que la différence de pression a atteint sa valeur maximum depuis plus de 200 ms .

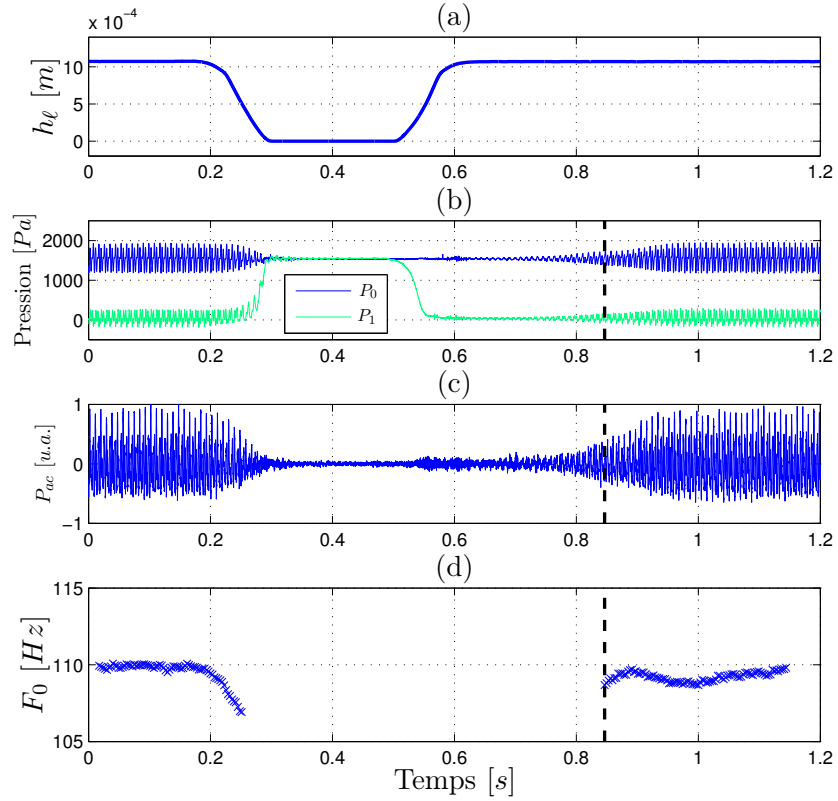


FIGURE 7.3 – Exemple de mesures réalisées pour $L_p = 10 \text{ cm}$. (a) : Aperture h_l des lèvres mécaniques. (b) : Pressions P_0 et P_1 mesurées respectivement en amont des maquettes de larynx et des lèvres. (c) : Pression acoustique p_{ac} mesurée à 20 cm au dessus de la maquette de larynx. (d) : Fréquence fondamentale F_0 du signal p_{ac} estimée à l'aide de l'algorithme Yin [dCK02]. Les droites verticales représentent l'estimation de l'instant correspondant à l'apparition des oscillations. [Ecouter le signal \$p_{ac}\$ \(son 13\)](#)

L'estimation précise de t_{on} est basée sur l'utilisation de l'algorithme Yin [dCK02]. L'algorithme, qui repose sur l'auto-corrélation du signal, fournit une valeur de la fréquence fondamentale F_0 uniquement lorsque la périodicité du signal dépasse un certain seuil. L'instant t_{on} est ici choisi comme étant celui de la première valeur détectée par l'algorithme après l'ouverture des lèvres mécaniques (cf. figure 7.3). Pour cet exemple, la valeur du VOT est $t_{on} - t_{ouv} = 350 \text{ ms}$, ce qui est largement supérieur aux valeurs de VOT relevées chez un locuteur humain, dont la durée moyenne est de l'ordre de 100 ms pour les plosives aspirées, 150 ms au maximum pour les amérindiens Navajos [CL99].

La fréquence fondamentale moyenne F_{0Moy} est également déterminée comme la moyenne des valeurs de F_0 entre les instants t_{on} et 700 ms après l'ouverture des lèvres mécaniques.

7.2.3 Résultats

La figure 7.4 présente les évolutions du VOT et de la fréquence fondamentale en fonction de la longueur du résonateur aval.

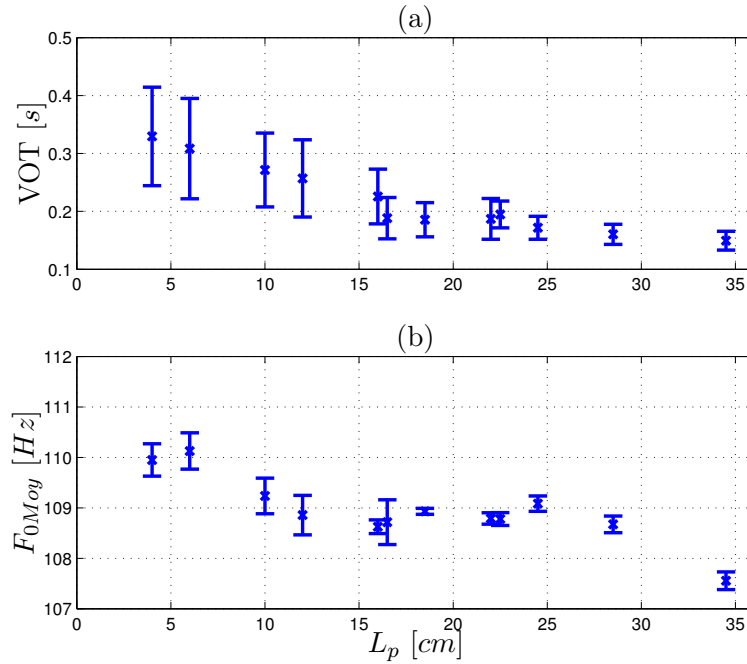


FIGURE 7.4 – (a) : Valeurs du VOT mesurées en fonction de la longueur L_p du tube en plexiglas inclus dans le résonateur aval. (b) : Fréquences fondamentales moyennes F_{0Moy} estimées en fonction de L_p . Les croix représentent la valeur moyenne sur 3 mesures comprenant chacune 10 répétitions. Les barres verticales représentent l'écart-type correspondant.

Ces résultats montrent que le VOT diminue avec l'augmentation de la longueur du résonateur aval. La fréquence fondamentale moyenne des oscillations suit cette même tendance, même si l'évolution n'est pas monotone. Il est également intéressant de noter que l'écart-type sur les valeurs de VOT mesurées, et sur les valeurs de F_{0Moy} dans une moindre mesure, diminue avec l'augmentation de L_p .

7.3 Simulations numériques

L'objectif de cette section est de vérifier que les modèles théoriques décrits au chapitre 2 permettent de reproduire les phénomènes observés sur la maquette de l'appareil phonatoire décrites en section 7.2.

Paramètres mécaniques	$m = 0,5 \text{ g}$, $k = 215 \text{ N.m}^{-1}$, $\xi = 0,025$, $k_c/k = 0,8$ $\eta_k = 8 \cdot 10^6 \text{ m}^{-2}$, $\eta_r = 5 \cdot 10^7 \text{ m}^{-2}$ $\xi_c = 1$, $\alpha_{kc} = 0$, $\alpha_{\eta kc} = 0$,
Paramètres géométriques	$L_g = 12 \text{ mm}$, $\ell_g = 12 \text{ mm}$, $x_1 - x_0 = x_3 - x_2 = 1 \text{ mm}$, $y_0 = 0,38 \text{ mm}$ $h_0 = h_3 = 16,8 \text{ mm}$

TABLE 7.1 – Paramètres du modèle à deux masses (cf. sections 2.2 et 2.3) utilisés pour simuler les mesures réalisées sur maquette lors de séquences voyelle - plosive non voisée - voyelle, pour différentes longueurs de résonateur aval.

7.3.1 Réalisation des simulations numériques

7.3.1.1 Implémentation du calcul du débit glottique

Lorsque la pression supra-glottique (ou intra-orale) ne peut pas être considérée comme nulle, le calcul de U_g est réalisé en appliquant le théorème de Bernoulli entre un point situé en amont des cordes vocales et un point situé devant les lèvres à l'extérieur du conduit vocal (équations 6.5 p. 121 et 6.9 p. 123). Cette méthode de calcul n'est pas compatible avec la prise en compte du couplage acoustique qui se traduit par la détermination de U_g à partir des ondes progressives p_{Mam}^+ et p_1^- incidentes à la glotte (équation 2.59).

La démarche adoptée est de réaliser ce calcul en deux étapes. Dans un premier temps, le débit U_g puis la pression P_{io} sont déterminés à partir des équations 6.5 et 6.9 sans considérer l'effet de la propagation acoustique. Le débit U_g est ensuite recalculé à partir de la pression trans-glottique $P_{sub} - P_{io}$ ainsi obtenue et à partir des ondes p_{Mam}^+ et p_1^- , en utilisant l'équation 2.59 (p. 48). Dans cette étude, l'expansion de la cavité supra-glottique (cf. section 6.4.1) n'est pas pris en compte dans le calcul de U_g .

7.3.1.2 Choix des paramètres

La géométrie des résonateurs amont et aval est relevée puis discrétisée en utilisant des tronçons cylindriques de longueur 5 mm . Avec une célérité acoustique c égale à 346 m.s^{-1} , la fréquence d'échantillonnage Fe obtenue est de 69262 Hz (cf. section 2.4). La géométrie discrétisée des résonateurs et les impédances d'entrée correspondantes sont données en annexe, figure D.2. La valeur minimum de la première fréquence de résonance acoustique, atteinte pour $L_p = 34,5 \text{ cm}$, est d'environ 150 Hz .

Les valeurs des paramètres mécaniques et géométriques du modèle à deux masses sont présentés dans le tableau 7.1. La géométrie et les dimensions de la constriction labiale correspondent à celles présentées en section 6.3.2.

Les paramètres d'entrée de la simulation sont la pression sous-glottique P_{sub} et l'aperture h_ℓ de la constriction labiale. Le signal P_{sub} est obtenu en filtrant passe-bas le signal de pression P_0 issu de la mesure de référence (figure 7.3). Ce filtrage est réalisé au moyen d'un filtre de Butterworth d'ordre 2 et de fréquence de coupure 30 Hz. Le signal h_ℓ est directement issu de la mesure de référence présentée en figure 7.3. Une interpolation polynomiale est réalisée pour adapter ces deux signaux à la fréquence d'échantillonnage de la simulation.

7.3.2 Méthode d'analyse des simulations

La figure 7.5 présente les résultats de la simulation réalisée pour L_p . La comparaison des évolutions de la fréquence fondamentale F_0 montre que, à la fermeture des lèvres pour $t = 250\text{ ms}$, les oscillations s'arrêtent au même instant tandis que, après l'ouverture des lèvres, l'absence de bruit d'écoulement fait que les oscillations simulées sont détectées plus tôt, pour $t = 0,7\text{ s}$. Dans le cas de la simulation, cet effet est compensé en déterminant l'instant t_{on} à partir des variations temporelles de F_0 . L'instant t_{on} correspond alors à l'instant où la dérivée temporelle dF_0/dt devient inférieure à 40 Hz.s^{-1} . Pour l'exemple présenté en figure 7.5, l'instant t_{on} est estimé à 805 ms .

La fréquence fondamentale moyenne F_{0Moy} est déterminée comme la moyenne des valeurs de F_0 entre l'instant t_{on} et la fin de la simulation, pour $t = 1,2\text{ s}$. Comme pour les mesures, les valeurs de VOT et de fréquence fondamentale F_{0Moy} sont relevées pour chaque longueur L_p .

7.3.3 Résultats

La figure 7.5.d montre que, lors de la fermeture des lèvres pour $200 \leq t \leq 280\text{ ms}$, la simulation permet d'observer une diminution de la valeur de F_0 , également observée expérimentalement. Lors de la reprise des oscillations, la simulation montre une forte augmentation de la valeur de F_0 ce qui correspond à la tendance générale observée sur les mesures, bien que ce phénomène soit moins marqué expérimentalement.

Les résultats pour l'ensemble des longueurs L_p testées sont présentés en figure 7.6. Ces résultats montrent, pour la simulation, une tendance à la diminution de la valeur du VOT lorsque L_p augmente. Une augmentation locale apparaît pour les valeurs de L_p comprises entre 16 et 22 cm inclus. En dehors de cet intervalle, l'erreur de prédiction reste inférieure à 10 %. L'évolution de F_{0Moy} simulée présente des différences importantes avec celle mesurée. La figure 7.6.b montre que F_{0Moy} atteint un maximum pour $L_p = 16\text{ cm}$, alors que dans le cas de la mesure, F_{0Moy} diminue progressivement avec L_p .

L'analyse détaillée des simulations réalisées montre que pour de faibles valeurs de L_p , les oscillations se produisent sans aucun contact entre les cordes vocales opposées (oscillations

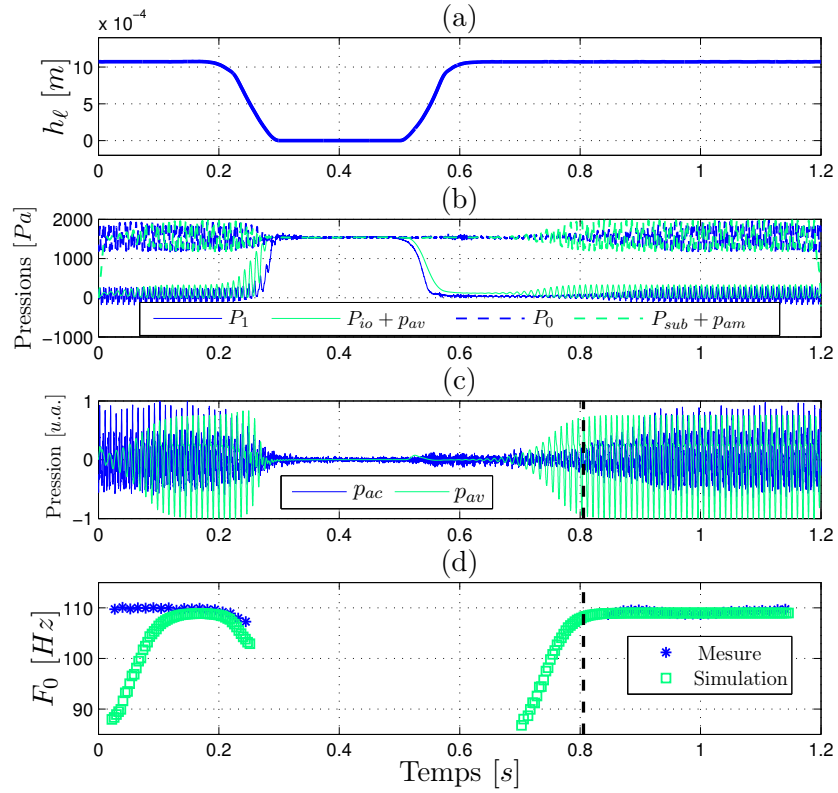


FIGURE 7.5 – Exemple de la simulation numérique réalisée pour $L_p = 10$ cm. (a) : Aperture h_ℓ des lèvres mécaniques mesurée, paramètre d'entrée de la simulation. (b) : Pressions mesurées, P_0 et P_1 , et simulées $P_{sub} + p_{am}$ et $P_{io} + p_{av}$ respectivement en amont des maquettes de larynx et des lèvres. (c) : Pression acoustique mesurée (p_{ac}) et pression acoustique simulée (p_{av}) en aval de la glotte. (d) : Fréquences fondamentales F_0 des signaux de pression mesuré (p_{ac}) et simulé (p_{av}), estimées à l'aide de l'algorithme Yin [dCK02]. Les droites verticales représentent l'estimation de l'instant t_{on} correspondant à l'apparition des oscillations pour la simulation.

[Ecouter le signal simulé \$p_{av}\$ \(son 12\)](#)

[Ecouter le signal mesuré \$p_{ac}\$ \(son 13\)](#)

correspondant au régime 2, cf. section 5.3.2). L'amplitude des oscillations augmente progressivement avec la longueur du résonateur aval jusqu'à l'apparition de collision au niveau de la deuxième masse pour $L_p = 16$ cm. La collision se produit également au niveau de la première masse à partir de $L_p = 18,5$ cm. Ce phénomène est associé au régime 1.

7.4 Discussion et conclusion

7.4.1 Discussion

Lien entre la pression de seuil et le Voice Onset Time

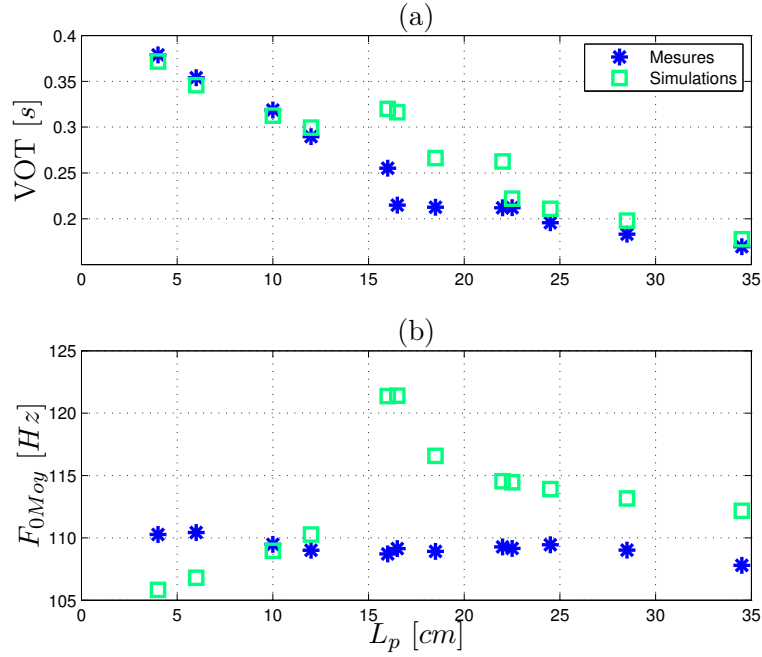


FIGURE 7.6 – (a) : Valeurs de VOT mesurées et simulées en fonction de la longueur L_p du tube en plexiglas inclus dans le résonateur aval. (b) : Fréquence fondamentale moyenne F_{0Moy} mesurée et simulée lors de la reprise des oscillations après l’ouverture, en fonction de L_p .

La différence maximale entre les différentes valeurs de VOT mesurées est de l’ordre de 200 ms, ce qui est très supérieure à la période, de l’ordre de 50 ms, pendant laquelle la pression P_1 chute, suite à l’ouverture des lèvres (cf. figure 7.3.b). Cette observation témoigne du fait que le retard à bifurcation n’est pas constant et qu’il varie de façon significative avec la longueur du résonateur aval.

Les résultats présentés en figure 7.4 montrent que l’hypothèse de départ est vérifiée : la diminution de la première fréquence de résonance acoustique a pour effet d’abaisser la pression de seuil P_{on} [Rut07, LLH⁺12], ce qui dans le cas des plosives se traduit par une réduction du retard à la bifurcation caractérisant l’apparition des oscillations. Les mesures réalisées (figure 7.4) montrent également que cette variation de la longueur du résonateur aval est accompagnée par une diminution de la fréquence fondamentale F_0 , ce qui est cohérent avec les études menées sur la pression de seuil [Rut07, LLH⁺12].

Validation des simulations numériques

Les simulations effectuées prédisent avec un bon accord les variations de VOT mesurés expérimentalement bien que la différence de méthode employée pour estimer l’instant t_{on} ne permette pas de valider une comparaison quantitative (figure 7.6.a). L’écart entre les valeurs de VOT mesurées et simulées augmente lorsque la collision entre les cordes vocales opposées se produit uniquement au niveau de la deuxième masse.

Interprétations phonétiques

Les résultats obtenus suggèrent que l'interaction entre les cordes vocales et le conduit vocal est potentiellement à l'origine de l'augmentation du VOT moyen lorsque le lieu où l'occlusion est réalisé est déplacé des lèvres vers le voile du palet [Kla75, Zue76, CH88, Byr93, Yao09]. À l'inverse, les résultats de ce chapitre ne permettent pas d'expliquer les tendances relevées par Fischer-Jørgensen [FJ72] et Klatt [Kla75] qui mentionnent des valeurs de VOT plus importantes lorsque la plosive est suivie par une voyelle fermée en comparaison avec le cas où celle-ci est suivie par une voyelle ouverte. Une explication possible est que ce phénomène est lié à la fréquence fondamentale intrinsèque des voyelles (cf. section 1.3.2). Les voyelles fermées sont produites avec une tension appliquée aux cordes vocales plus importante, ce qui entraîne une augmentation de P_{on} (cf. section 1.2.3) et donc une augmentation du VOT.

Limite du modèle mécanique

Les variations de la fréquence fondamentale moyenne F_{0moy} simulées (figure 7.6), peuvent être analysées en fonction de la présence de collision. F_{0moy} augmente avec L_p dans le cas où la collision ne se produit pas et diminue dans le cas contraire. Dans le premier cas, l'augmentation de F_{0Moy} observée, qui n'est pas observée expérimentalement, peut s'expliquer par l'utilisation d'un modèle mécanique incluant une expression non-linéaire de la force élastique (équation 2.19). En effet, la diminution de F_1 a pour effet d'augmenter l'amplitude des oscillations, ce qui engendre une augmentation de la raideur équivalente et donc de la fréquence fondamentale F_0 . Dans le cas où la collision se produit, l'amplitude des oscillations est limitée, ce qui réduit l'augmentation de F_0 . Cette interprétation s'appuie également sur le fait que cette augmentation n'est pas observée pour des simulations réalisées en l'absence de non-linéarités dans l'expression de la force élastique et en l'absence de collision ([DPBL14]). La modification du comportement induit par l'apparition du phénomène de collision peut être interprétée comme une limite des modèles à 2 dimensions utilisés dans ces travaux. Sur la maquette, les deux tubes de latex représentant les cordes vocales sont initialement en contact au niveau de leurs extrémités. Lors de l'auto-oscillation, la surface impliquée dans la collision est donc partielle et progressive tandis que dans le cas de la simulation, la collision (lorsqu'elle a lieu) se produit systématiquement sur la totalité de la longueur des cordes.

7.4.2 Conclusion

Les mesures et les simulations effectuées dans ce chapitre mettent en évidence le lien existant dans le cas des plosives bilabiales entre la pression de seuil des oscillations et le retard à la bifurcation de Hopf. Dans le cas où l'effet du couplage acoustique est faible, c'est-à-dire lorsque la fréquence fondamentale reste inférieure à la fréquence F_1 de la première résonance acoustique, une diminution de la pression de seuil des oscillations due à une augmentation de la longueur du résonateur aval, entraîne également une diminution du VOT. Une faible pression de seuil favorise donc une apparition rapide des oscillations auto-entretenues.

Les simulations numériques mettent en avant le fait que les modèles théoriques utilisés permettent de reproduire les variations de VOT en fonction de F_1 , ainsi que l'évolution de la fréquence fondamentale lorsque la collision entre l'une des deux masses du modèle se produit.

Dans ce chapitre, les simulations et les mesures ont été réalisées dans le cas de plosives bilabiales. Une perspective intéressante serait de reproduire ces expériences en utilisant un résonateur de longueur constante et en faisant varier progressivement le lieu du conduit vocal où l'occlusion est réalisée.

Les simulations réalisées font également apparaître une limite dans la modélisation de la collision. À l'instar de Pelorson et al. [PHvH⁺94] qui proposent d'introduire un paramètre supplémentaire pour tenir compte de la fermeture partielle des cordes vocales au niveau aérodynamique, la prise en compte de la collision partielle au niveau du comportement mécanique des cordes vocales pourrait permettre d'améliorer le réalisme des simulations.

Conclusion et perspectives

Les travaux réalisés dans cette thèse ont pour objectif une meilleure compréhension des phénomènes physiques gouvernant la production de la parole, notamment dans l’optique de contribuer au développement d’un outil de synthèse par modèle physique. Les études effectuées portent sur les troubles de la production de la parole liés à un défaut de contrôle du voisement. Le premier cas étudié est l’apparition de sauts abruptes dans l’évolution de la fréquence fondamentale lors de la phonation. Le deuxième cas concerne le contrôle temporel du voisement lors de la production de consonnes plosives.

Modélisation physique de la production de la parole

Bilan

Les simulations numériques réalisées pour les deux études menées dans cette thèse reposent sur les mêmes modèles théoriques. Les écoulements d’air glottique et labial sont décrits par l’équation de Bernoulli corrigée pour la viscosité. Le modèle mécanique utilisé pour décrire le mouvement des cordes vocales est basé sur le modèle à deux masses de Lous [LHVH98]. Afin d’étendre la validité de la modélisation à des oscillations de grande amplitude, des termes non-linéaires sont ajoutés à l’expression des forces visqueuse et élastique. La propagation acoustique est modélisée à l’aide de guides d’ondes numériques, basés sur l’approximation d’ondes planes progressives. La géométrie des résonateurs acoustiques amont et aval est alors approximée par des agrégations de tronçons cylindriques. Les effets des pertes viscothermiques au niveau des parois lors de la propagation et des réflexions aux extrémités des résonateurs sont pris en compte en utilisant des filtres numériques.

Conclusions

Malgré sa simplicité, le modèle à deux masses de Lous permet de reproduire les principales caractéristiques dynamiques de l’auto-oscillation des cordes vocales. Le degré de complexité des modèles utilisés est suffisant pour reproduire les bifurcations correspondant à l’apparition et à l’extinction des oscillations entretenues et aux transitions de mécanisme laryngé. Les modèles utilisés sont similaires à ceux employés dans la littérature pour décrire le comportement de la source laryngée en voix parlée. Cependant, les simulations reproduisant des cas de production pour lesquels le contrôle de la phonation est faible, font apparaître des limitations importantes. En effet, les résultats obtenus montrent que lorsque la fréquence fondamentale de la production est proche d’une fréquence de résonance acoustique, l’influence des interactions acoustiques est largement sur-estimée. La fréquence fondamentale de l’auto-oscillation est alors imposée par les résonateurs acoustiques et l’amplitude des signaux de pression acoustique simulés devient équivalente à celle de la pression sous-glottique. Ce comportement est amplifié lorsque l’amplitude de la pression sous-glottique dépasse la gamme des valeurs correspondant à la voix parlée.

Les autres limites de ce modèle simplifié viennent du fait que la dimension antero-postérieure de la constriction glottique est constante. Cette approximation a des conséquences importantes sur la façon dont est simulée la collision au niveau de l'écoulement et au niveau du mouvement des cordes vocales.

Perspectives

Les conclusions mettent en avant la limite principale des modèles théoriques utilisés, à savoir la sur-estimation de l'effet du couplage acoustique. Deux solutions sont envisageables pour résoudre ce problème. La première est de prendre en compte l'effet de la vibration des parois du conduit vocal sur la propagation acoustique. Les effets de ce phénomène dissipatif sont localisés en basse fréquence. La vibration des parois peut réduire significativement l'amplitude du premier pic de l'impédance d'entrée du conduit vocal. La deuxième possibilité est de modifier la modélisation du couplage entre l'aérodynamique et l'acoustique au niveau du larynx. Suivre les propositions de Laje et Mindlin [LM05], qui considèrent que le débit glottique est une source acoustique sphérique et que l'approximation des ondes planes dans le conduit vocal n'est valide qu'à partir d'une certaine distance de la glotte, pourrait constituer une perspective.

Dans ces travaux, les modèles utilisés dans les parties aérodynamique, mécanique ou acoustique sont relativement simples. La complexité vient essentiellement des interactions mécanique-aérodynamique d'une part et aérodynamique-acoustique d'autre part. Dans ces conditions, il est particulièrement difficile d'assurer la stabilité numérique de l'ensemble du système. Ce constat suggère l'utilisation d'un formalisme qui garantirait la cohérence du comportement énergétique [BS05] ou bien du bilan de puissance de l'ensemble du système [LH14].

Sauts de fréquence fondamentale

Bilan

Des observations acoustiques et électroglottographiques ont été réalisées pour les trois voyelles cardinales produites par un locuteur sourd profond post-linguistique et implanté. Dans certains cas, pour les voyelles fermées (/i/ et /u/) produites avec une forte intensité, un phénomène de saut de fréquence fondamentale est observé. La valeur de la fréquence fondamentale passe d'environ 500 à 400 Hz dans un intervalle de temps de 30 ms . Les signaux électroglottographiques montrent que ces sauts de fréquence fondamentale sont associés à une transition du mécanisme laryngé II vers le mécanisme I. Une maquette auto-oscillante de l'appareil phonatoire a été utilisée pour imiter ce phénomène. Des transitions entre deux régimes, accompagnées par un saut de la fréquence fondamentale, ont été observées en faisant varier trois paramètres de contrôle différents : en diminuant la raideur des répliques de corde vocale, en augmentant le niveau de la pression en aval des cordes vocales artificielles et en augmentant la longueur du résonateur aval. Ces expériences sont reproduites par simulations numériques. La comparaison entre les prédictions numériques et les données expérimentales permet d'évaluer les modèles utilisés. Afin d'étudier les stratégies motrices expliquant l'apparition de transitions de mécanisme laryngé, les simulations sont réalisées en faisant varier la

raideur des cordes vocales, la pression sous-glottique et l'aire glottique initiale, pour les trois voyelles cardinales. Les résultats sont analysés sous la forme de cartographies des différents régimes dans un espace formé par ces trois paramètres de contrôle.

Conclusions

Les mesures sur maquette et les simulations numériques indiquent que les transitions de mécanisme laryngé sont les manifestations d'une bifurcation du système laryngé. Les simulations réalisées font apparaître deux régimes d'oscillations distincts. Le motif de vibration des deux masses pour ces deux régimes présente des similitudes importantes avec le mouvement des cordes vocales observé pour les productions en mécanismes laryngés I et II. Le régime 1 se caractérise par des faibles valeurs de fréquence fondamentale et de quotient ouvert et par une collision au niveau de deux masses tandis que le régime 2 est caractérisé par une fréquence fondamentale et un quotient ouvert plus importants et par l'absence de collision au niveau de la masse située à l'extrémité amont de la glotte.

Les transitions entre les régimes 1 et 2 sont accompagnées par un saut de fréquence fondamentale. Les modèles utilisés permettent de reproduire l'amplitude des sauts de fréquence observés sur la maquette avec une précision de l'ordre de 25 %. Les transitions observées sur maquette et simulées font apparaître un cycle d'hystérésis qui témoigne de la coexistence des deux régimes pour une région de paramètres de contrôle donnée. Les simulations reproduisent ce phénomène bien que la taille de la région pour laquelle les deux régimes coexistent soit sur-estimée lorsque le couplage avec le résonateur acoustique est pris en compte.

Les cartographies des différents régimes d'oscillation simulés montrent que, par rapport au régime 1, le régime 2 est associé à des valeurs importantes de la raideur des cordes vocales, de l'aire glottique initiale, et à de faibles valeurs de pression sous-glottique. Ces analyses montrent également que, dans la région de paramètres de contrôle pour laquelle les régimes 1 et 2 coexistent, les oscillations font apparaître des comportements quasi-périodiques parfois chaotiques.

Vers une rééducation des troubles de la production chez les sourds

Pour un locuteur sourd, l'apparition de sauts de fréquence fondamentale est due à l'utilisation involontaire du mécanisme laryngé II en voix parlée. Dans le cas étudié, ce mécanisme apparaît pour les voyelles fermées (/i/ et /u/), pour lesquelles la première fréquence de résonance acoustique est relativement basse, produites avec une forte intensité. Les études numériques et expérimentales suggèrent que le couplage acoustique n'est pas responsable de l'apparition du mécanisme II. L'origine de ce trouble peut être liée à la position haute de la langue, qui induit une augmentation de la tension des cordes vocales [Hon83]. Proposer une solution de rééducation spécifique nécessiterait de déterminer la stratégie musculaire utilisée par les locuteurs normo-entendants pour maintenir une fréquence fondamentale relativement faible lors de la production de voyelle fermées.

Perspectives

Au niveau des observations *in vivo*, il serait bien sûr intéressant de reproduire ces me-

sures en incluant d'avantage d'informations expérimentales, telles que des mesures de pression sous-glottique et de l'impédance acoustique du conduit vocal durant la phonation. Ces informations supplémentaires permettraient de préciser les interprétations sur les stratégies motrices responsables de l'apparition de ces phénomènes. À notre connaissance, la reproduction sur maquette de sauts de fréquence fondamentale associés à des transitions de mécanisme laryngé n'apparaît pas dans la littérature. Le comportement dynamique du système laryngé pourrait cependant être étudié plus en détails, par exemple en mesurant l'évolution de la fréquence fondamentale lors de sauts de fréquence provoqués par une augmentation puis une diminution de la raideur des cordes vocales, pour différentes longueurs de résonateurs amont ou aval.

Plosives bilabiales

Bilan

Des mesures acoustiques et aérodynamiques *in vivo* ont été réalisées durant la production de la séquence /apa/. Les mouvements d'ouverture et de fermeture des lèvres sont analysés à l'aide d'une caméra ultra-rapide à une cadence de 500 Hz. Des mesures visant à reproduire une séquence voyelle - plosive non voisée - voyelle sont ensuite réalisées sur une maquette de l'appareil phonatoire. Ces mesures permettent de valider les modèles aérodynamiques utilisés pour décrire l'écoulement d'air dans les lèvres et l'effet de l'expansion de la cavité supraglottique. L'effet de l'interaction acoustique entre le conduit vocal et les cordes vocales est étudié par le biais de mesures sur maquette et de simulations numériques, en modifiant la longueur du résonateur aval. Les données correspondantes sont analysées en déterminant la valeur du délai entre le relâchement de l'occlusion et l'établissement du voisement (*Voice Onset Time* ou VOT).

Conclusions

Les simulations réalisées montrent que la production d'une séquence voyelle - plosive non voisée - voyelle ne requiert aucun ajustement laryngé. L'équilibrage de la pression de part et d'autre du larynx entraîne une augmentation passive de l'aire glottique moyenne (abduction des cordes vocales) et un arrêt de l'auto-oscillation. L'augmentation du volume de la cavité supra-glottique, sous l'effet de la pression intra-orale, est un phénomène essentiel à la description des interactions aérodynamiques. Ce phénomène est responsable du maintien du voisement pendant environ 40 ms après la fermeture des lèvres.

D'autre part, les simulations et les mesures réalisées sur maquette indiquent que la longueur du résonateur aval influe considérablement sur la valeur du VOT. Ces résultats pourraient expliquer les variations observées chez les locuteurs humains : pour les consonnes plosives, la valeur moyenne du VOT augmente lorsque le lieu où l'occlusion est réalisée est déplacé des lèvres vers le voile du palais, quel que soit le contexte phonétique ou le langage employé. Ce comportement s'explique par l'influence du couplage acoustique qui favorise l'apparition des oscillations, notamment lorsque la première fréquence de résonance du conduit vocal est proche et supérieure à la fréquence fondamentale de la phonation.

Vers une rééducation des troubles de la production chez les sourds

Les résultats suggèrent que l'expansion de la cavité supra-glottique joue un rôle important dans le maintien du voisement pendant toute la durée de l'occlusion lors de la production d'une séquence voyelle - plosive bilabiale voisée - voyelle. Chez les personnes sourdes, les troubles de la production liés à une substitution de plosive voisée par son homologue non voisée (et inversement) pourraient être corrigés en utilisant un outil de monitoring du volume de la cavité formée entre le larynx et l'occlusion.

Perspectives

Des réalisations complémentaires de simulations de séquences voyelle - plosive voisée - voyelle permettraient d'analyser plus précisément l'influence de l'expansion de la cavité supraglottique sur le maintien du voisement. Il serait notamment intéressant de déterminer si l'expansion passive, due à la pression intra-orale, est suffisante pour produire une plosive voisée ou bien si une augmentation active du volume de la cavité supra-glottique, induite par des ajustements articulatoires comme l'abaissement de la mâchoire ou le larynx, est nécessaire. D'autre part, une étude supplémentaire incluant des mesures et des simulations avec une longueur de conduit vocal fixe et avec différentes positions de l'occlusion permettrait d'appuyer les conclusions sur l'influence du lieu d'articulation sur le VOT. Ces travaux montrent que le couplage acoustique avec les résonateurs a une influence significative sur la valeur du VOT. Il serait intéressant d'approfondir cette étude en s'intéressant à l'influence du type de voyelle ou bien à l'influence de phénomènes physiques tels que les pertes dues à la vibration des parois du conduit vocal et les effets viscothermiques lors de la propagation acoustique.

Couches limites et séparation de l'écoulement

A.1 Introduction

L'objectif de cette annexe est de fournir un cadre théorique afin d'expliquer le phénomène de séparation de l'écoulement. Les hypothèses conduisant à la théorie des couches limites sont présentées puis les notions relatives à cette théorie sont utilisées pour expliquer les causes de la séparation de l'écoulement dans le cas de la constriction glottique. Les hypothèses formulées en section 2.2.1 sont adoptées.

A.2 Hypothèses

Le modèle à deux masses utilisé considère que la dimension de la constriction glottique est constante dans la direction antero-postérieure, correspondante au vecteur \vec{e}_z . De plus, la longueur L_g associée à cette direction est grande devant l'épaisseur des cordes vocales ℓ_g selon l'axe x et devant la hauteur de la $h(x)$ selon l'axe y . Le fait que pour la glotte $L_g \gg \ell_g$ et $L_g \gg h(x) \forall x$ permet de justifier l'hypothèse d'un écoulement plan, pour lequel la composante v_z est nulle et la vitesse ne dépend pas de z .

Écoulement unidimensionnel

Pour un écoulement plan, l'équation de continuité pour un fluide localement incompressible (équation 2.1), qui s'écrit $\partial v_x / \partial x + \partial v_y / \partial y = 0$, impose que $v_x / \ell_g \sim v_y / h(x)$ et donc que $v_y / v_x \sim h(x) / \ell_g$. De même, l'ordre de grandeur des variations de vitesse selon les directions principale et transversale est lié aux dimensions caractéristiques de l'écoulement, ainsi $\partial^2 v_x / \partial x^2 \sim v_x / \ell_g^2$ et $\partial^2 v_x / \partial y^2 \sim v_x / h(x)^2$. En considérant que $h(x) \ll \ell_g \forall x$, l'équation de continuité implique donc que $v_y \ll v_x$ et $\partial^2 v_x / \partial x^2 \ll \partial^2 v_x / \partial y^2$.

A.3 Théorie des couches limites

Équations de Prandtl

Pour un écoulement quasi-stationnaire et unidimensionnel, les équations de Navier-Stokes (2.1 et 2.2), projetées sur les axes x et y , pour un écoulement instationnaire, se réduisent aux équations de Prandtl,

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0, \quad (\text{A.1})$$

$$v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial x} + \nu \frac{\partial^2 v_x}{\partial y^2}, \quad (\text{A.2})$$

$$\frac{\partial P}{\partial y} = 0. \quad (\text{A.3})$$

Pour un écoulement dans une conduite, ce système d'équations est associé à la condition de non-glissement au niveau de la paroi $v_x|_{y=0} = 0$. Cette condition est à l'origine du gradient longitudinal de vitesse. Il est supposé qu'à une certaine distance de la paroi, v_x ne dépend plus de y . Le modèle de Prandtl considère que les effets visqueux sont localisés au niveau des parois.

Notion de couche limite

Le nombre de Reynolds Re mesure le rapport entre les effets d'inertie et les effet visqueux. Nous utilisons l'expression proposée par Vilain [Vil02], tel que $Re = v h^2 / (\nu \ell)$. Dans le cas d'un écoulement de Stokes ($Re < 1$), les forces d'inertie sont négligeables. Pour l'écoulement d'air au travers la glotte, l'ordre de grandeur du nombre de Reynolds est $Re \approx 500$. L'écoulement peut être considéré comme laminaire : les effets d'inertie sont prépondérants devant ceux dus à la viscosité du fluide. Cependant, les effets visqueux ne peuvent être négligés dans une zone proche des parois de la constriction. La description aérodynamique repose alors sur la séparation de l'écoulement en deux régions. L'écoulement principal, distant des parois, est supposé unidimensionnel et le fluide est supposé parfait (effets de viscosité négligeables). L'écoulement pariétal, également appelé couche limite, est supposé bi-dimensionnel et visqueux. Le comportement de la couche limite est régi par les équations de Prandtl (équations A.1, A.2 et A.3). Celui de l'écoulement principal est décrit par l'équation d'Euler, tel que :

$$v_x \frac{\partial v_x}{\partial x} = -\frac{1}{\rho} \frac{\partial P}{\partial x}. \quad (\text{A.4})$$

La cohérence du modèle est assurée par la condition d'égalité de la vitesse v_x à l'interface entre ces deux régions.

Couche mince

Étant donné que le gradient de pression transverse (équation A.3) est négligeable, l'unique variation de pression est orientée suivant la direction de l'écoulement et dépend du

profil de la constriction. L'épaisseur caractéristique de la couche limite δ est donc régie par deux processus dont les effets s'opposent. D'une part, la viscosité, par le biais de la diffusion de la quantité de mouvement des parois vers l'écoulement principal, tend à augmenter δ . D'autre part, le transfert d'énergie cinétique de l'écoulement principal vers la couche limite tend à diminuer l'épaisseur δ .

En supposant l'égalité des temps caractéristiques visqueux (δ^2/ν) et convectif (ℓ/v), il est possible d'estimer l'épaisseur de la couche limite [Vil02] :

$$\frac{\delta}{h} = \sqrt{\frac{\nu \ell}{v h^2}} = \frac{1}{\sqrt{Re}}. \quad (\text{A.5})$$

Le rapport $\frac{\delta}{h}$ est donc de l'ordre de 10^{-2} , la couche limite peut donc être considérée comme mince.

A.4 Séparation de l'écoulement

Pour un écoulement incompressible et quasi-stationnaire, l'équation de continuité se traduit par la conservation de la masse. Une particule de fluide, qui traverse une constriction convergente est donc accélérée ($\frac{\partial v_x}{\partial x} > 0$), tandis qu'une particule traversant une constriction divergente subit une décélération. En aval de la constriction, la décélération de l'écoulement engendre un gradient de pression qui s'oppose à l'écoulement. Lorsque l'énergie cinétique de l'écoulement principal n'est plus suffisante pour entraîner les particules de la couche limite, l'écoulement se sépare de la paroi pour former d'un jet libre. La condition de séparation de l'écoulement est exprimée à l'aide de la contrainte pariétale de cisaillement τ_0 , telle que :

$$\tau_0 = \mu \left. \frac{\partial v_x}{\partial y} \right|_{y=0} \leq 0. \quad (\text{A.6})$$

En aval du point de séparation, l'interaction entre la zone pariétale dans laquelle le fluide est au repos et le jet engendre l'apparition d'un comportement turbulent, qui se traduit par une forte dissipation de l'énergie cinétique du jet, sans modification significative de la pression. Nous considérons un jet quasi-stationnaire, associé à une pression constante (la pression en aval de la constriction est égale à la pression au point de séparation de l'écoulement).

Calibration de la station EVA

B.1 Introduction

Ce travail a été réalisé en collaboration avec Christophe Savariaux, ingénieur de recherche et Xavier Laval, ingénieur d'étude au département Parole et Cognition du Gipsa-lab.

B.1.1 Motivations

Ce document a pour objectif de présenter les résultats des mesures de calibration de la station EVA2, réalisées les 14 décembre 2011 et 4 Janvier 2012 au département Parole et Cognition du Gipsa-Lab. La station EVA (Evaluation Vocale Assistée) permet de réaliser différentes mesures aéro-acoustiques pendant la production de parole.

La motivation principale de ce travail vient des problèmes de synchronisation observés lors de mesures faisant intervenir la station EVA et une caméra vidéo associée à une autre carte d'acquisition numérique. La caméra est utilisée pour extraire les paramètres labiaux du sujet durant la production de plosives bilabiales [Mai11]. Malgré la synchronisation ponctuelle entre la station EVA et la vidéo, réalisée au début de chaque session de mesure, un retard apparaît sur les signaux enregistrés par la station EVA par rapport au signal vidéo. L'ordre de grandeur de ce retard est de quelques millisecondes, ce qui n'est pas important devant la durée de certains phénomènes intervenant dans la production de parole (voisement, fricatives) mais ce qui pose un problème pour étudier des événements plus courts

tels que les plosives.

Deux types de retards différents ont été identifiés :

- un retard initialement nul qui augmente progressivement au cours de la mesure, pour l'ensemble des signaux de la station EVA, par rapport à la vidéo,
- un retard de quelques millisecondes sur la voie 3, utilisée pour les capteurs de pression (intra-orale ou sous-glottique)

Une calibration statique pour les capteurs de pression et de débit oral a également été réalisée.

B.1.2 Description de la station EVA

La station EVA2 a été développée à l'université d'Aix en Provence (UR 261 , laboratoire Parole et Langages), avec la collaboration du CHU Timone de Marseille [TG95, TGG99, GYR⁺06]. La station EVA dispose d'un capteur de pression acoustique et de plusieurs capteurs aérodynamiques :

- un microphone à condensateur, de directivité hypercardioïde (AKG, modèle C419)
- deux capteurs de pression différentielle,
 - un capteur de faible pression, couramment utilisé pour mesurer la pression intra-orale,
 - un capteur de forte pression, utilisé pour mesurer la pression sous-glottique,

- deux capteurs de débit (pneumotachographes à grille),
 - un capteur de débit oral, placé à l'extrémité d'un masque, appliqué sur le bas du visage pendant les mesures,
 - un capteur de débit nasal, composé de deux orifices reliés au sujet par deux tubes en plastique souple insérés dans les narines.

Il est important de noter pour la suite que les capteurs de pression sont connectés à la voie d'acquisition numéro 3 de la station. Cette voie peut également être utilisée pour acquérir d'autres signaux physiologiques (EGG, pression oesophagienne, ...).

La figure B.1 présente la pièce à main de la station EVA, sur laquelle est disposée l'ensemble des capteurs.

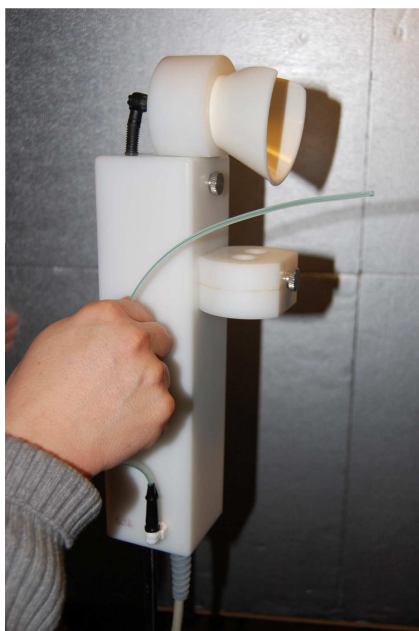


FIGURE B.1 – Pièce à main de la station EVA.

B.1.3 Matériel et capteurs de référence

Les mesures de calibration ont été réalisées au département Parole et Cognition du Gipsa-Lab, sur une maquette du système phonatoire humain. Le matériel expérimental utilisé comme référence, est celui de la salle aéro-acoustique 2. L'acquisition des données se fait par un conditionneur de signaux, une carte d'acquisition National Instruments (PCI-MIO-16XE-10) et le logiciel Labview. Le capteur de pression utilisé est un capteur de pression différentielle piezo-résistif (Endevco, 8507c-2). Il a été préalablement calibré avec un manomètre à eau. Le capteur de débit de référence est un débitmètre massique (TSI, 4043).

Ces mesures de calibration s'organisent en 3 manipulations différentes :

- une calibration statique des capteurs de pression et de débit oral,
- une mesure de la dérive temporelle de la carte d'acquisition de la station EVA,
- une calibration dynamique des capteurs de pression, connectés à la voie numéro 3 de la station.

La station EVA existe également en version portable, qui présente les mêmes fonctionnalités. La station EVA portable fait également l'objet de mesures de calibration statique et d'estimation de la dérive temporelle.

B.2 Calibration statique

B.2.1 Protocole expérimental

Ces mesures sont réalisées sur une maquette du système phonatoire. Un débit d'air

constant est créé dans un cylindre en métal, au moyen d'une surpression dans le réservoir de pression disposé en amont de l'écoulement. Le réservoir de pression est alimenté par un compresseur, une vanne permet de contrôler manuellement la pression qui alimente le réservoir. La pression mesurée à l'intérieur du cylindre est acheminée vers le capteur de pression d'EVA et le capteur de référence par le biais d'un connecteur « en forme de T » et de deux tubes en plastique de diamètre $0,5\text{ cm}$ et de longueur 30 cm . Le débitmètre de référence est connecté à la sortie du tube en métal. En aval de l'écoulement, un tube en plastique de diamètre $1,5\text{ cm}$ et de longueur 180 cm , fait le lien avec le capteur de débit oral de la station EVA.

La pression et donc le débit d'air sont augmentés progressivement au cours de chaque mesure. Les valeurs de pression et de débit présentées dans cette section sont obtenues en moyennant les signaux sur les paliers où les amplitudes sont constantes. Cette manipulation est répétée quatre fois pour couvrir l'ensemble des capteurs et échelles de mesures de la station. Le tableau B.1 présente les valeurs maximales de deux échelles de mesures pour les deux capteurs de pression et le capteur de débit oral.

Capteurs	Petite échelle	Grande échelle
Faible pression	20 hPa	100 hPa
Forte pression	40 hPa	200 hPa
Débit	2 L.s^{-1}	10 L.s^{-1}

TABLE B.1 – Valeurs maximales des différentes échelles d'acquisition pour les deux capteurs de pression et pour le capteur de débit oral des stations EVA et EVA portable.

B.2.2 Station EVA

B.2.2.1 Pression

La figure B.2 présente les mesures de calibration des capteurs de pression de la station EVA. Les courbes représentant les valeurs de pression données par la station EVA sont des droites de pente légèrement supérieure à la droite de pente unité. Pour les trois configurations testées ici, les capteurs de pression sont linéaires jusqu'à 15 hPa .

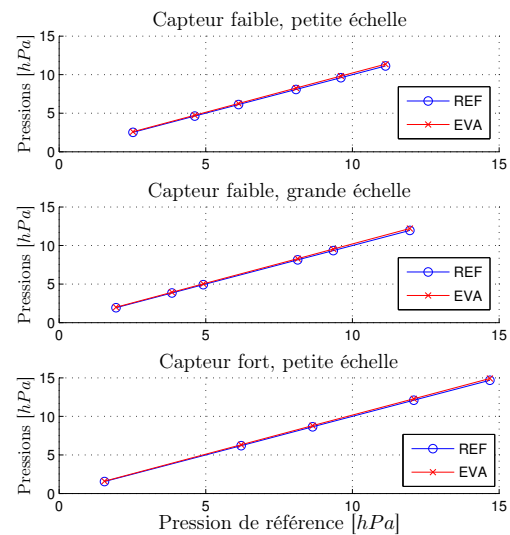


FIGURE B.2 – Mesures de calibration statique des capteurs de pression de la station EVA, pour deux échelles de mesure différentes.

Les données expérimentales sont modélisées par une droite dont la pente est estimée par régression linéaire. Le tableau B.2 donne les écarts relatifs des pentes estimées par rapport à la pente unité pour chaque configuration, ainsi que les coefficients de calibration correspondants. Les valeurs de pression données par la station EVA sont sur-estimées, l'écart entre les valeurs de référence et celles de la station EVA reste inférieur à 2 %.

Capteurs	Faible pression		Forte pression
Échelles	petite	grande	petite
Écart	+1,9%	+1,9%	+1,3%
Coefficients	0,9815	0,9811	0,9868

TABLE B.2 – Écarts relatifs entre les pentes obtenues à partir des valeurs de pression et des coefficients de calibration correspondants pour la station EVA.

B.2.2.2 Débit

La figure B.3 montre les résultats obtenus pour les deux échelles de mesure du capteur de débit oral. Dans ce cas, la linéarité du capteur n'est pas aussi satisfaisante que dans le cas précédent, en particulier pour des débits faibles (inférieurs à $0,3 \text{ L.s}^{-1}$). D'autre part, un écart considérable est observé entre les valeurs de débit données par la station EVA et le débitmètre pour la petite échelle de mesure.

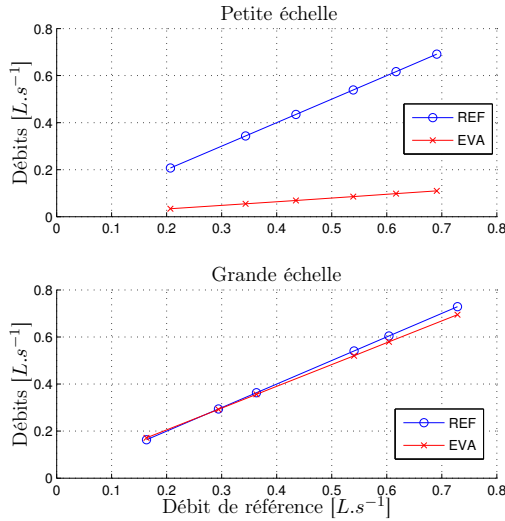


FIGURE B.3 – Mesures de calibration statique du capteur de débit oral de la station EVA pour deux échelles de mesure différentes.

De la même manière que précédemment, la pente des courbes est obtenue par régression

linéaire, le tableau B.3 donne les écarts entre les pentes ainsi que les coefficients de calibration correspondants. Il apparaît que le débit donné par la station EVA est sous-estimé de 84 % pour la petite échelle et de 7 % pour la grande échelle de mesure.

Échelles	petite	grande
Écart	−84,3%	−7,4%
Coefficients	6,3705	1,0798

TABLE B.3 – Écarts relatifs entre les pentes obtenues à partir des valeurs de débit et coefficients de calibration correspondants pour la station EVA.

B.2.3 Station EVA portable

La station EVA portable a également fait l'objet de mesures de calibration statique utilisant le même protocole.

B.2.3.1 Pression

La figure B.4 présente les résultats de ces mesures pour les différentes échelles et capteurs de pression. La linéarité des deux capteurs de pression de la station EVA portable est satisfaisante. Dans le cas de la grande échelle de mesure pour le capteur de forte pression, la pression donnée par la station EVA portable est considérablement plus élevée que la pression de référence. Dans les autres cas, la pression est légèrement sur-estimée.

Ces écarts sont quantifiés dans le tableau B.4. Pour la grande échelle de mesure, sur le capteur de forte pression, la pression est sur-estimée de 400 %.

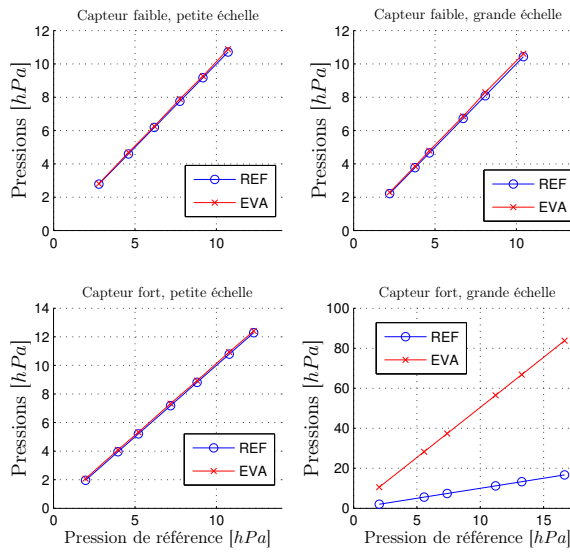


FIGURE B.4 – Mesures de calibration statique des capteurs de pression de la station EVA portable pour deux échelles de mesure différentes.

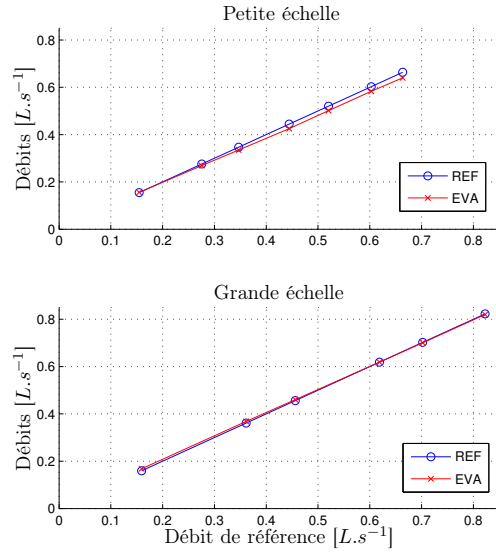


FIGURE B.5 – Mesure de calibration statique du capteur de débit oral de la station EVA portable pour deux échelles de mesure différentes.

Capteurs	Faible pression		Forte pression	
Échelles	petite	grande	petite	grande
Écarts	+1,4%	1,4%	+0,1%	+400%
Coefficients	0,9863	0,9860	0,9990	0,1999

TABLE B.4 – Écarts relatifs entre les pentes obtenues à partir des valeurs de pression et coefficients de calibration correspondants pour la station EVA portable.

Les écarts entre les pentes obtenues par régression linéaire sont donnés par le tableau B.5. L'écart maximum est de 4.6 % pour la petite échelle du capteur de débit oral.

Échelles	petite	grande
Écarts	-4,6%	-1,9%
Coefficients	1,0482	1,0198

TABLE B.5 – Écarts relatifs entre les pentes obtenues à partir des valeurs de débit et coefficients de calibration correspondants pour la station EVA portable.

B.2.3.2 Débit

La figure B.5 présente les mesures de calibration statique pour les deux échelles du capteur de débit oral. Le capteur de débit oral de la station EVA portable a le même comportement que celui de la station EVA, le débit est légèrement sous-estimé et le capteur n'est pas complètement linéaire pour les faibles débits.

B.2.4 Discussion

La limitation imposée par le compresseur fait que les valeurs de pression et de débit utilisées pour la calibration ne couvrent pas entièrement la plage de mesure de la station EVA, en particulier dans le cas des grandes échelles

de mesure. La calibration réalisée ici est cependant suffisante dans le cas d'une utilisation de la station pour la production de parole. Ces mesures montrent que la tendance générale est que les stations EVA et EVA portable sur-estiment les valeurs de pression et sous-estiment les débit oraux. Ce résultat général a déjà été mis en évidence lors d'un stage-ingénieur [Hau03].

La tendance observée pour les débits des stations EVA et EVA portable peut venir du protocole expérimental. En effet, la longueur importante du tube qui relie le débitmètre de référence au capteur de la station EVA peut expliquer la présence de pertes de charge non-négligeables. Ce phénomène pourrait également expliquer les problèmes de linéarité observés pour de faibles débits. Ces écarts sont acceptables sauf dans les cas particuliers de la petite échelle de débit pour la station EVA et pour la grande échelle sur le capteur de forte pression pour la station EVA portable.

Une explication possible pour les faibles signaux de débit donnés par la station EVA dans le cas de la petite échelle serait le dysfonctionnement ou la panne d'un composant électronique, qui contribue à un gain analogique. Les valeurs de pression 5 fois plus importantes données par la station EVA portable (pour le capteur fort et de la grande échelle), ne peuvent provenir que d'un problème de gain numérique quelque part dans la chaîne d'acquisition.

un autre signal audio synchrone avec la vidéo. Au bout de 6 minutes d'enregistrement, ce retard atteint environ 4 ms.

B.3.1 Protocole expérimental

La manipulation consiste en un enregistrement audio simultané sur trois appareils différents :

- la station EVA (ou EVA portable),
- la carte d'acquisition National Instruments (NI),
- un enregistreur numérique portable Marantz PMD670 (PMD).

Les fréquences d'échantillonnage sont de 25 kHz pour la station EVA et la carte NI et de 44,1 kHz pour le PMD. Le PMD est utilisé comme horloge de référence. Ce choix s'explique par le fait qu'il est couramment utilisé au département Parole et Cognition pour les expériences faisant intervenir de l'audio et de la vidéo, la dérive temporelle entre le PMD et la caméra utilisée est négligeable et vérifiée régulièrement. La synchronisation est réalisée à l'aide d'un signal acoustique impulsionnel produit par « clap » à proximité des trois microphones et répété toutes les minutes environ.

B.3.2 Station EVA

La figure B.6 montre l'évolution du retard de la station EVA et de la carte NI en fonction du temps du PMD. L'évolution du retard est linéaire pour les deux appareils. La dérive temporelle est définie comme le rapport de la différence entre le temps mesuré et le temps de référence, pour un même événement acoustique, sur le temps de référence de l'événement. Elle est estimée au moyen

B.3 Dérive temporelle

L'objectif de cette mesure est de déterminer la déviation temporelle de la carte d'acquisition numérique de la station EVA. Des mesures utilisant la station et une caméra vidéo [Mai11] ont mis en évidence la présence d'un retard du signal audio de la station EVA, qui augmente progressivement, par rapport à

d'une régression linéaire. Le tableau B.6 présente la dérive temporelle estimée ainsi que la correction qu'il faut appliquer à la fréquence d'échantillonnage de 25 kHz à la lecture des signaux audio. D'après ces mesures, la dérive temporelle est plus importante sur la carte NI que pour EVA. Pour la station EVA, la dérive temporelle est de $1,31.10^{-3}\%$, ce qui sur 6 minutes équivaut à un retard de $4,7\text{ ms}$, il s'agit donc bien du phénomène précédemment identifié.

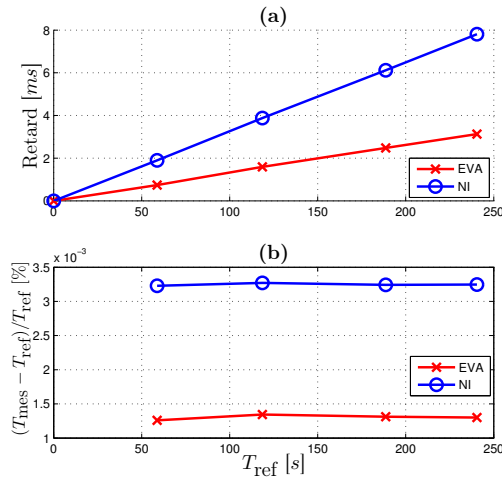


FIGURE B.6 – Mesures de la dérive temporelle de la station EVA et de la carte NI par rapport au PMD. (a) : Évolution du retard mesuré en fonction du temps de référence. (b) : Dérive temporelle en fonction du temps de référence.

Appareils	EVA	NI
Dérive temporelle	$1,31.10^{-3}\%$	$3,25.10^{-3}\%$
Correction sur F_e	$+0,3272\text{ Hz}$	$+0,8118\text{ Hz}$

TABLE B.6 – Dérives temporelles de la station EVA portable et de la carte NI et corrections de la fréquence d'échantillonnage à prendre en compte à la lecture des fichiers numériques.

B.3.3 Station EVA portable

Cette mesure de déviation temporelle a aussi été réalisée pour la station EVA portable. La carte NI a également été testée ce qui fournit deux réalisations différentes de cette mesure.

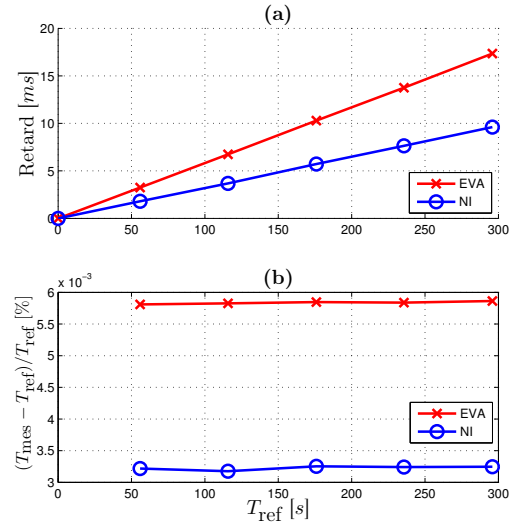


FIGURE B.7 – Mesures de la dérive temporelle de la station EVA portable et de la carte NI par rapport au PMD. (a) : Évolution du retard mesuré en fonction du temps de référence. (b) : Dérive temporelle en fonction du temps de référence.

Les résultats obtenus sont présentés dans la figure B.7 et le tableau B.7 donne les valeurs de la dérive temporelle et des corrections de la fréquence d'échantillonnage associées.

Appareils	EVA	NI
Dérive temporelle	$5,86.10^{-3}\%$	$3,25.10^{-3}\%$
Correction sur F_e	$+1,4648\text{ Hz}$	$+0,8130\text{ Hz}$

TABLE B.7 – Dérives temporelles de la station EVA portable et de la carte NI et corrections de la fréquence d'échantillonnage à prendre en compte à la lecture des fichiers numériques.

Il est intéressant de noter que pour les deux réalisations, l'estimation de la dérive temporelle de la carte NI est similaire (Tableau B.6). Cette mesure est donc bien répétable. La précision de l'horloge interne donnée par le constructeur pour la carte d'acquisition NI est $10^{-2}\%$. Les valeurs estimées de la dérive temporelle de la carte NI et des stations EVA et EVA portable restent inférieures à cette limite. Les phénomènes de déviation temporelle entre deux appareils sont courants. Dans le cas d'une acquisition sur des durées importantes, il est essentiel de synchroniser les horloges internes des différents appareils.

Remarque : Les tableaux B.6 et B.7 donnent les corrections de la fréquence d'échantillonnage pour $F_e = 25 \text{ kHz}$, pour le signal audio des stations EVA et EVA portable. Les autres signaux de la station EVA sont échantillonnés à $F_e = 6,25 \text{ kHz}$, soit quatre fois moins que pour l'audio. Dans la partie suivante, où l'on s'intéresse à la calibration dynamique des capteurs de pression de la station EVA, la fréquence d'échantillonnage pour la lecture des signaux de pression est choisie telle que $F_e = 25000,3272/4 \text{ Hz}$.

B.4 Calibration dynamique

L'objectif de cette manipulation est d'identifier et quantifier le retard qui apparaît sur les signaux de pression enregistrés sur la voie 3 de la station EVA.

B.4.1 Protocole expérimental

Ces mesures sont réalisées en utilisant le même dispositif que pour la calibration statique (c.f. section B.2.1). Dans cette manipulation, le débit d'air est modulé périodique-

ment par les fermetures et ouvertures successives d'une constriction située en aval du point de mesure. La fréquence du moteur pas-à-pas qui contrôle l'ouverture de la constriction varie au cours de la manipulation entre 1 et 5 Hz. Nous obtenons ainsi un signal de pression pseudo-périodique avec une composante continue, et dont la fréquence varie au cours du temps. Au début de chaque mesure, les deux signaux sont synchronisés grâce un signal audio sinusoïdal de courte durée (« bip ») enregistré simultanément sur les deux appareils. La correction de fréquence d'échantillonnage déterminée dans la section précédente a également été prise en compte.

Les signaux sont filtrés passe-bas, avec un filtre Butterworth d'ordre 5, dont la fréquence de coupure est fixé à 50 Hz. La figure B.8 montre l'allure des deux signaux de pression une fois que ces recalages temporels ont été effectués. Il apparaît que le signal issu de la station EVA est en retard de quelques millisecondes.

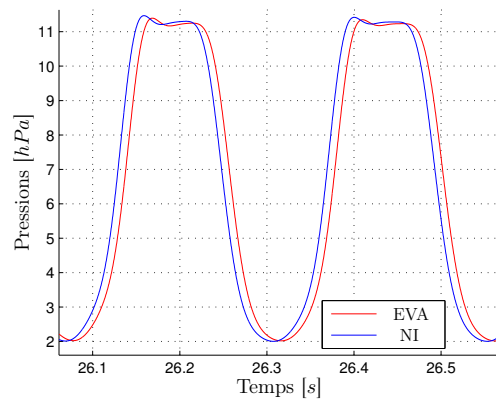


FIGURE B.8 – Exemple de deux périodes de signaux de pression obtenus après la synchronisation initiale et la correction de la fréquence d'échantillonnage

B.4.2 résultats pour la station EVA

L'objectif est de déterminer si l'origine du retard est due à un phénomène physique, lié à l'inertie de l'air dans le tube, ou bien à un problème d'ordre métrologique. Nous faisons l'hypothèse que si l'origine du retard est physique alors ce retard varie en fonction de vitesse de variation de la pression, et donc en fonction de la fréquence du moteur. Le décalage temporel entre les deux signaux est estimé par intercorrélation sur des fenêtres glissantes de 25 secondes avec un taux de recouvrement de 75 %.

Les figures B.9 et B.10 présentent les résultats obtenus pour les deux échelles de mesure du capteur de faible pression et les figures B.11 et B.12 montrent les résultats pour le capteur de forte pression. Ces 4 figures sont organisées de la manière suivante :

Graphique (a) Estimation du retard de la voie 3 de la station EVA par intercorrélation sur des fenêtres glissantes de 25 s en fonction du temps.

Graphique (b) Fréquence d'oscillation du signal de pression en fonction du temps.

Graphique (c) Estimation du retard par intercorrélation en fonction de la fréquence d'oscillation du signal de pression.

Capteur/Échelle	Petite	Grande
Faible Pression	8,80 ms	8,64 ms
Forte Pression	8,64 ms	8,64 ms

TABLE B.8 – Estimation du retard de la voie 3 de la station EVA par intercorrélation entre les signaux de pression entiers

Les figures B.9 à B.12 montrent que le retard de la voie 3 de la station EVA est constant et similaire pour chacune des configurations de capteur et d'échelle. Les faibles variations sont

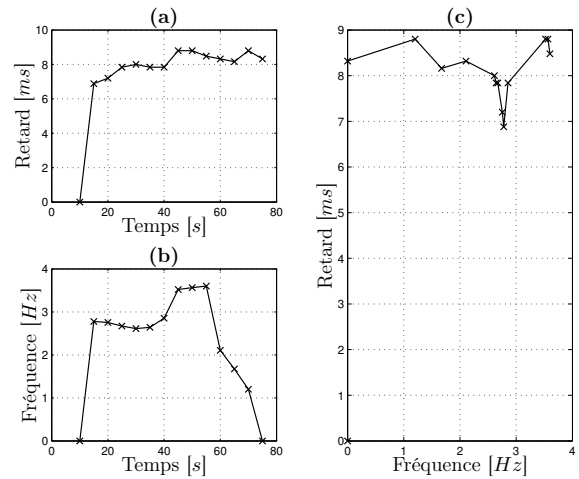


FIGURE B.9 – Capteur de faible pression, petite échelle de mesure

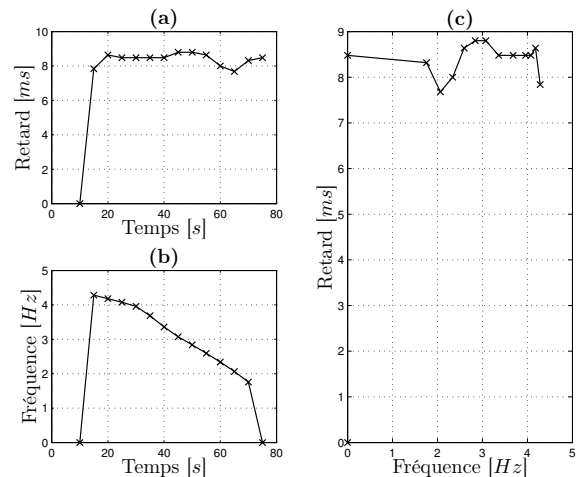


FIGURE B.10 – Capteur de faible pression, grande échelle de mesure

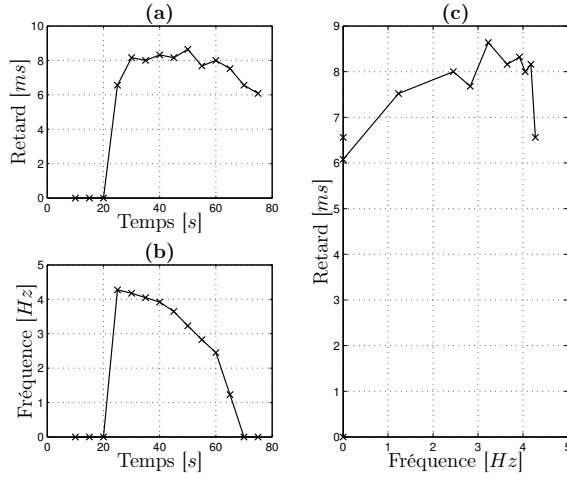


FIGURE B.11 – Capteur de forte pression, petite échelle de mesure

dues à la méthode d'estimation. L'interprétation est que l'origine de ce retard est d'ordre métrologique. Nous utilisons ensuite une intercorrélation sur l'ensemble des signaux de façon à obtenir une estimation plus précise de ce retard supposé constant. Le tableau B.8 donne les estimations de ces retards. La valeur du retard estimé dans le cas de la petite échelle de mesure pour le capteur de faible pression est la seule qui diffère des autres cas. Cette différence peut-être due à une erreur de mesure. Nous considérons que ce retard de $8,64 \text{ ms}$ est le même pour tous les capteurs et échelles de pression et qu'il est donc inhérent à la voie 3 de la station EVA.

B.5 Conclusion

Les stations EVA et EVA portable ont fait l'objet d'une calibration statique pour les capteurs de pression et de débit oral. Ces mesures ont mis en évidence deux problèmes notables. Le premier concerne le capteur de débit de la station EVA pour la petite échelle de mesure et le deuxième concerne le capteur de forte pression de la station EVA portable pour la grande échelle de mesure.

La dérive temporelle des cartes d'acquisition numériques des deux stations EVA a également été estimée. En prenant comme référence un enregistreur numérique portable, nous mesurons une dérive temporelle de $1,3 \cdot 10^{-3} \%$ pour la station EVA et de $5,9 \cdot 10^{-3} \%$ pour la station EVA portable. Un retard systématique de $8,64 \text{ ms}$ a été identifié sur la voie 3 de la station EVA qui est notamment utilisée pour les capteurs de pression intra-orale et sous-glottique.

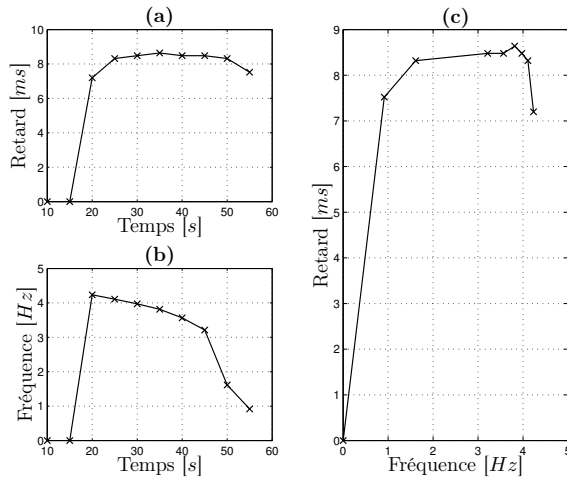


FIGURE B.12 – Capteur de forte pression, grande échelle de mesure

Mesures complémentaires de sauts de fréquence fondamentale

La figure C.1 présente deux exemples de sauts de fréquence observés en fin de production pour la voyelle /i/ produite par le sujet Y. Dans les deux cas, la fréquence fondamentale F_0 passe de 550 à 450 Hz et le quotient ouvert Q_o de 75 à 40 % durant un intervalle de temps d'environ 30 ms . Après le saut de fréquence, l'estimation des fréquences de résonance acoustique par prédiction linéaire donne $F_1 \simeq 800$ Hz et $F_2 \simeq 1400$ Hz pour l'exemple présenté à gauche sur la figure C.1 et $F_1 \simeq 800$ Hz et $F_2 \simeq 1600$ Hz pour l'exemple présenté à droite de la même figure. Ces valeurs correspondent approximativement à la voyelle ouverte /a/.

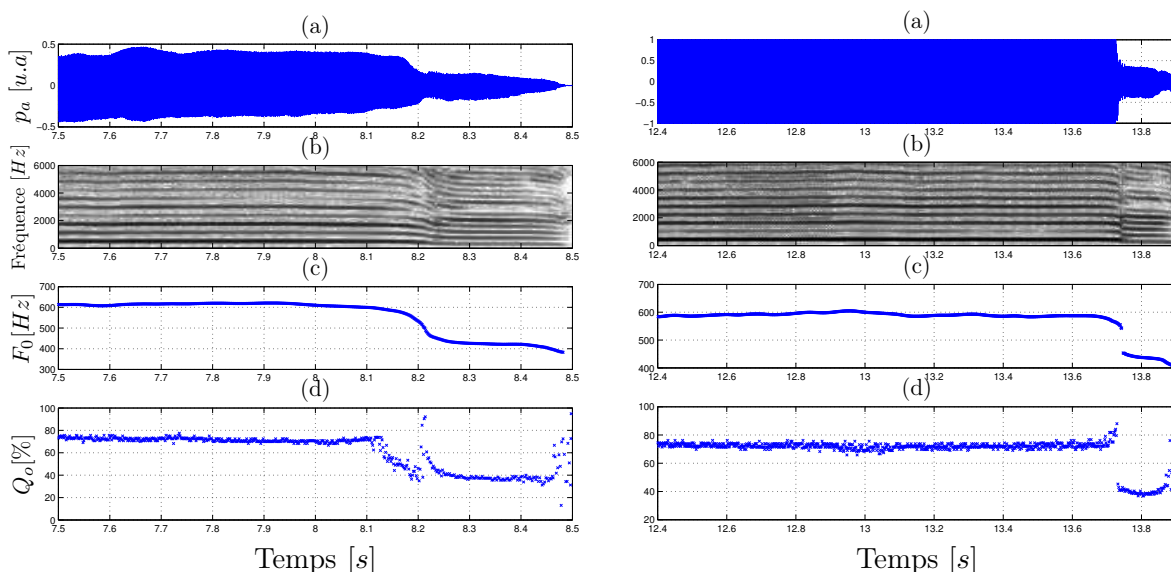


FIGURE C.1 – Deux exemples de sauts de fréquence mesurés pour la voyelle /i/ produite avec une intensité *crescendo* par le sujet Y. (a) : Signal de pression acoustique p_a . (b) : Spectrogramme de p_a calculé en utilisant des fenêtres de 10 ms . (c) : Fréquence fondamentale F_0 du signal p_a . (d) : Quotient ouvert Q_o extrait du signal EGG .

La figure C.2 présente un exemple de saut de fréquence fondamentale produit par un autre locuteur, appelé sujet C. Ces données sont issues des mesures réalisées durant le stage qui a précédé ce travail de thèse [Mai11]. Le sujet C est une femme d'environ 30 ans souffrant d'une surdité profonde intervenue avant l'acquisition du langage. Le sujet C a reçu une éducation basée sur l'oralisation.

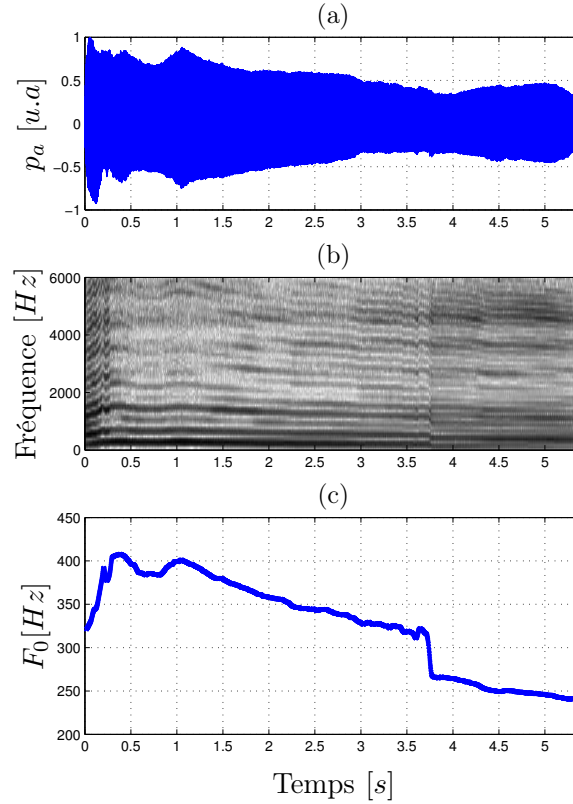


FIGURE C.2 – Exemple de saut de fréquence mesuré pour la voyelle /u/ produite avec une forte intensité par le sujet C. (a) : Signal de pression acoustique p_a . (b) : Spectrogramme de p_a calculé à l'aide de fenêtres de 10 ms. (c) : Fréquence fondamentale F_0 du signal p_a .

La figure C.2 montre un saut de fréquence fondamentale négatif, lors de la production d'une voyelle fermée, pour $t = 3,7$ s, alors que la valeur de F_0 décroît régulièrement avant et après l'événement. Ce saut s'accompagne d'un enrichissement spectral visible sur le spectrogramme. Pour ces mesures, la qualité des signaux *EGG* ne permet pas de déterminer l'évolution du quotient ouvert. Cependant, une augmentation de l'asymétrie du signal *EGG* est observée après le saut de fréquence fondamentale. Ces observations permettent de déduire que ce phénomène est dû à une transition de mécanisme laryngé dans le sens MI→MII. L'estimation des fréquences de résonance du conduit vocal par prédiction linéaire avec 38 pôles, donne pour cette mesure, une valeur de F_1 variant progressivement de 430 à 490 Hz et une valeur de F_2 relativement constante, environ égale à 830 Hz avant le saut de fréquence puis à 1100 Hz après. L'augmentation de la valeur de F_1 suggère qu'un mouvement d'ouverture progressif de la constriction est réalisé durant la production, grâce à un abaissement de la langue par exemple.

Résonateurs acoustiques

D.1 Reproduction par simulation des mesures réalisées sur maquette

Les figures D.1 et D.2 présentent respectivement la géométrie discrétisée et l'impédance d'entrée des résonateurs acoustiques amont et aval, utilisées en simulation pour reproduire les mesures décrites en sections 4.4 et 7.2.

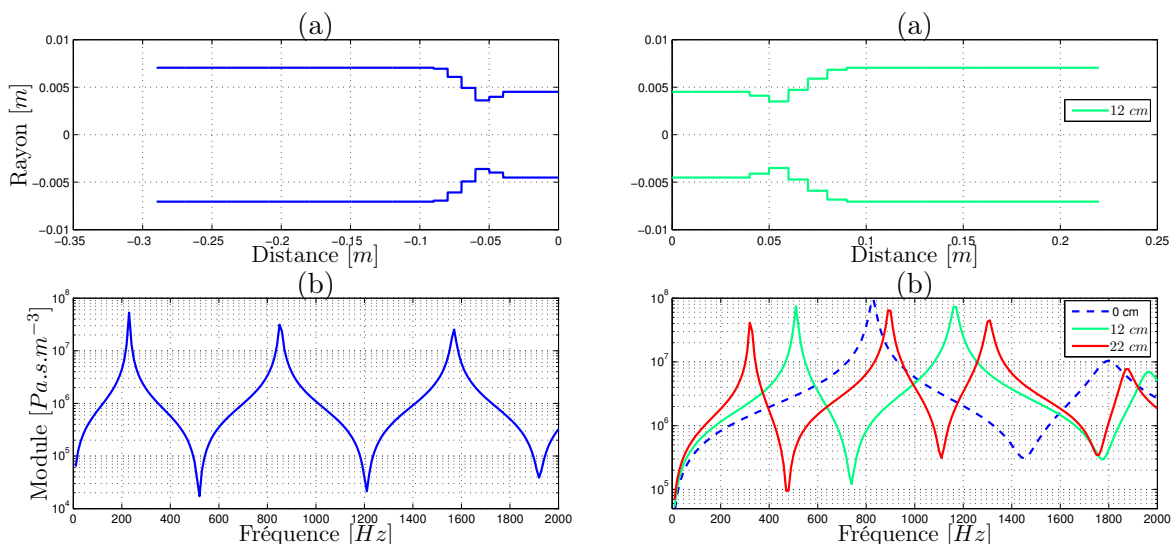


FIGURE D.1 – Colonne de gauche : résonateur amont (trachée). Colonne de droite : résonateur aval (conduit vocal) pour différentes longueurs L_p . (a) : Approximation géométrique, pour la simulation, des résonateurs acoustiques par des tronçons cylindriques de 1 cm de longueur. La distance nulle correspond à la position de la glotte. Le tube en plexiglas est connecté au point d'abscisse $x = 9,5$ cm. (b) : Impédance d'entrée des résonateurs calculée à partir de la géométrie discrétisée en utilisant l'équation 2.55.

D.2 Simulation de parole produite par un locuteur humain

La figure D.3 présente les résonateurs acoustiques utilisés pour modéliser la trachée et le conduit vocal dans les simulations effectuées au chapitre 5. La section du résonateur amont

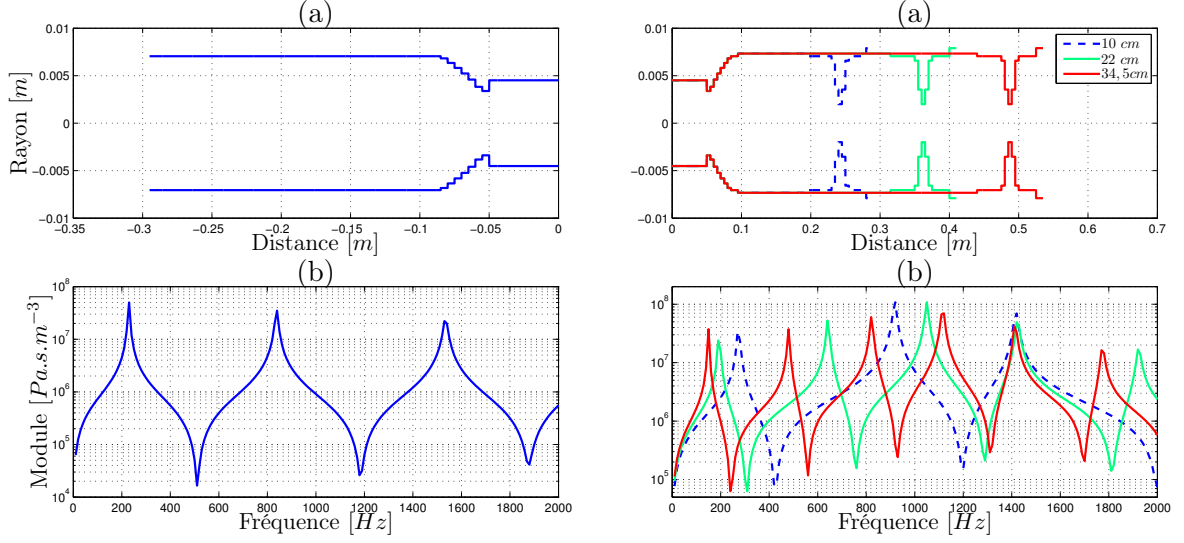


FIGURE D.2 – Colonne de gauche : résonateur amont (trachée). Colonne de droite : résonateur aval (conduit vocal) pour différentes longueurs L_p . (a) : Approximation géométrique, pour la simulation, des résonateurs acoustiques par des tronçons cylindriques de 5 mm de longueur. (b) : Impédance d'entrée des résonateurs calculée à partir de la géométrie discrétisée en utilisant l'équation 2.55.

est choisie de façon à être égale à la section de l'entrée de la glotte : $L_g \times h_0 = 3,6 \cdot 10^{-4} m^2$ et sa longueur est de 17,5 cm. Ce choix se justifie par la valeur de la première fréquence de résonance obtenue (520 Hz) qui est environ égale à la valeur moyenne mesurée par Arsi-kere et al. [ALLA13] pour les hommes.

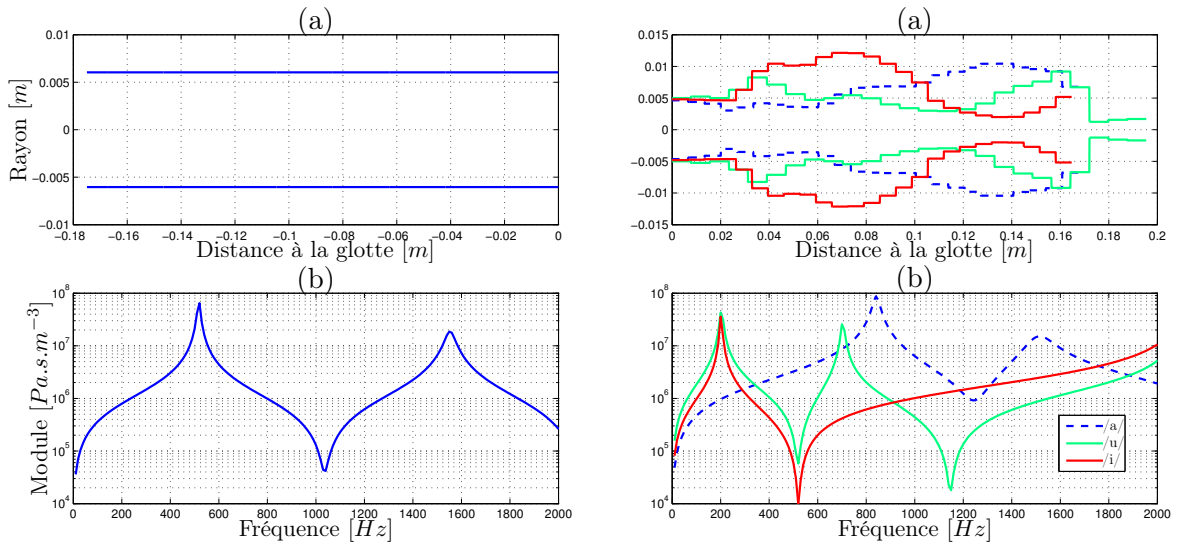


FIGURE D.3 – Colonne de gauche : résonateur amont (trachée). Colonne de droite : résonateur aval (conduit vocal) pour les trois voyelles cardinales. (a) : Approximation géométrique, pour la simulation, des résonateurs acoustiques par des tronçons cylindriques dont la longueur est comprise entre 6,6 et 7,8 mm . La distance nulle correspond à la position de la glotte. Le tube en plexiglas est connecté au point d'abscisse $x = 9,5$ cm . (b) : Impédance d'entrée des résonateurs calculée à partir de la géométrie discrétisée en utilisant l'équation 2.55.

Cartographies des régimes d'oscillation

La figure E.1 présente les cartographies des régimes d'oscillation simulés pour la voyelle /i/ avec une valeur de masse $m = 0,05 \text{ g}$. Les figures E.2 et E.3 montrent les résultats obtenus dans le cas où la masse est deux fois plus importante ($m = 0,1 \text{ g}$), lorsque le couplage acoustique n'est pas pris en compte et dans le cas de la voyelle /a/. Les paramètres acoustiques correspondants sont donnés pour les régimes 1 et 2 en figure E.4.

L'analyse de la figure E.1 montre que le comportement des simulations pour la voyelle /i/ est très proche de celui des simulations réalisées pour la voyelle /u/ (cf. figure 5.5). Ce constat peut s'interpréter par le fait que les fréquences des premières résonances des deux résonateurs utilisés sont très proches (situées à 200 Hz). La différence remarquable est que, le phénomène d'accrochage de F_0 sur la première fréquence de résonance du conduit vocal, pour $160 < F_0 < 200 \text{ Hz}$, qui se produit dans le cas de la voyelle /u/, pour $k = 50 \text{ N.m}^{-1}$ et $P_{sub} > 2700 \text{ Pa}$, n'est pas observé pour la voyelle /i/.

Lorsque le couplage acoustique n'est pas considéré, l'organisation des régimes d'oscillation obtenue avec une masse de $0,1 \text{ g}$ (figure E.2) est quasiment identique à celle obtenue pour une masse plus faible (figure 5.3). La comparaison entre les figures 5.3 et E.2 montre que, lorsque le couplage acoustique n'est pas considéré, l'influence d'une augmentation de la masse se traduit essentiellement par une diminution de la fréquence fondamentale F_0 des oscillations. L'analyse des figures 5.4 et E.3 met en avant le fait que, dans le cas de la voyelle /a/, la diminution des valeurs de F_0 modifie considérablement la localisation des régions de paramètres correspondants à des phénomènes d'accrochage de F_0 sur les fréquences de résonance acoustique.

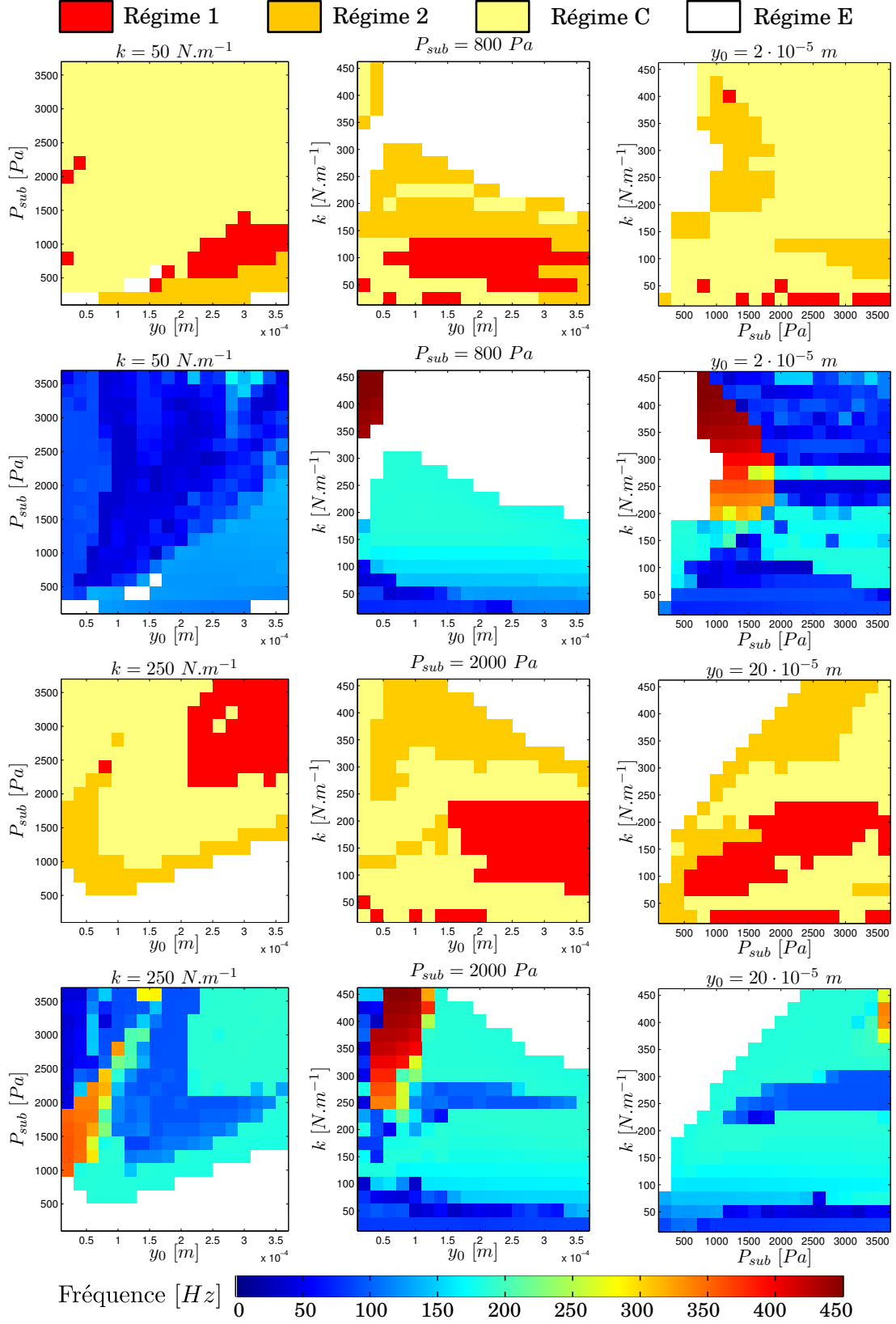


FIGURE E.1 – Cartographie des régimes d'oscillation simulés pour la voyelle /i/ et pour une masse de $m = 0,05 g$, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 .

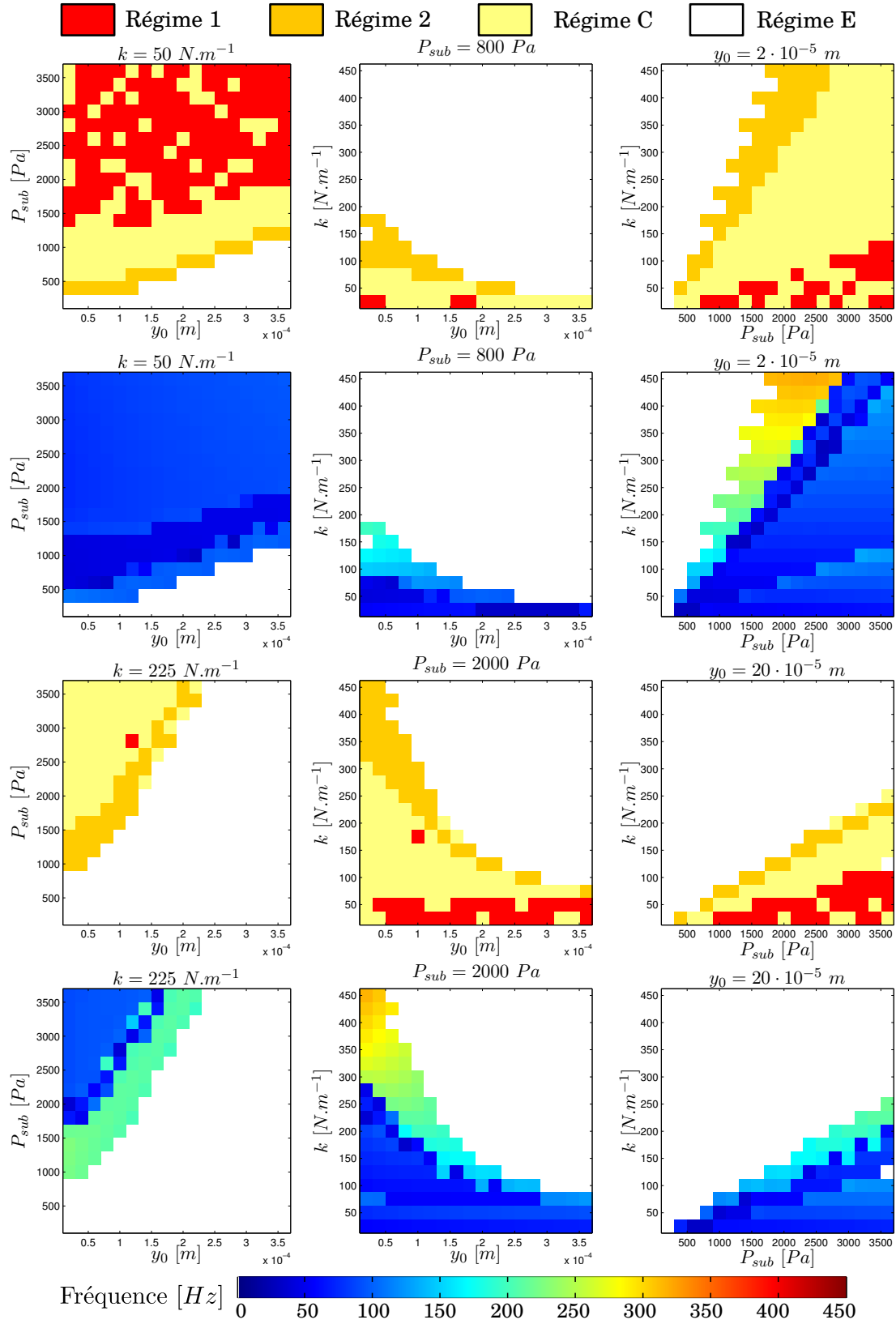


FIGURE E.2 – Cartographie des régimes d’oscillation simulés sans couplage acoustique et pour une masse de $m = 0,1 \text{ g}$, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 .

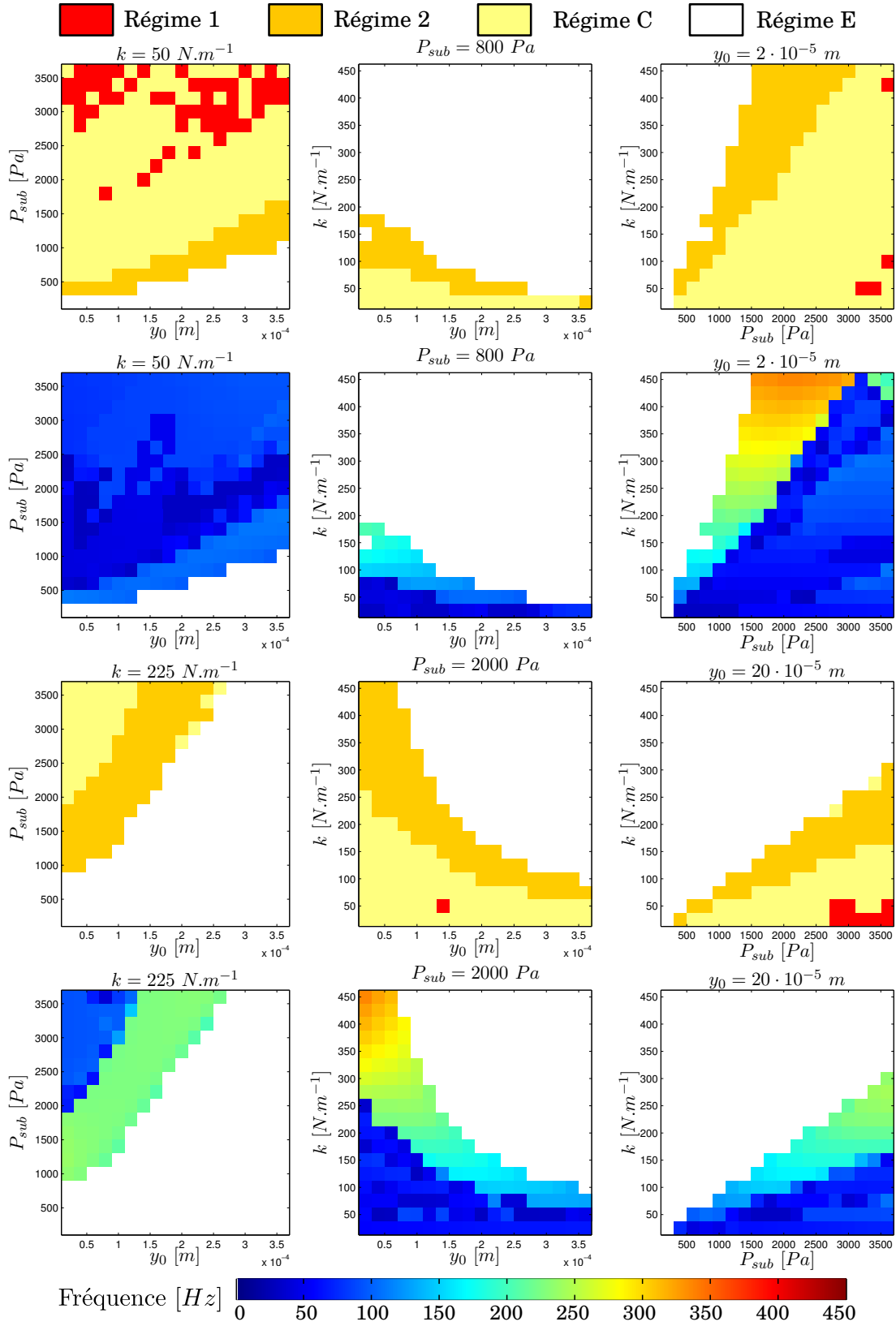


FIGURE E.3 – Cartographie des régimes d'oscillation simulés pour la voyelle /a/ et pour une masse de $m = 0,1 \text{ g}$, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 .

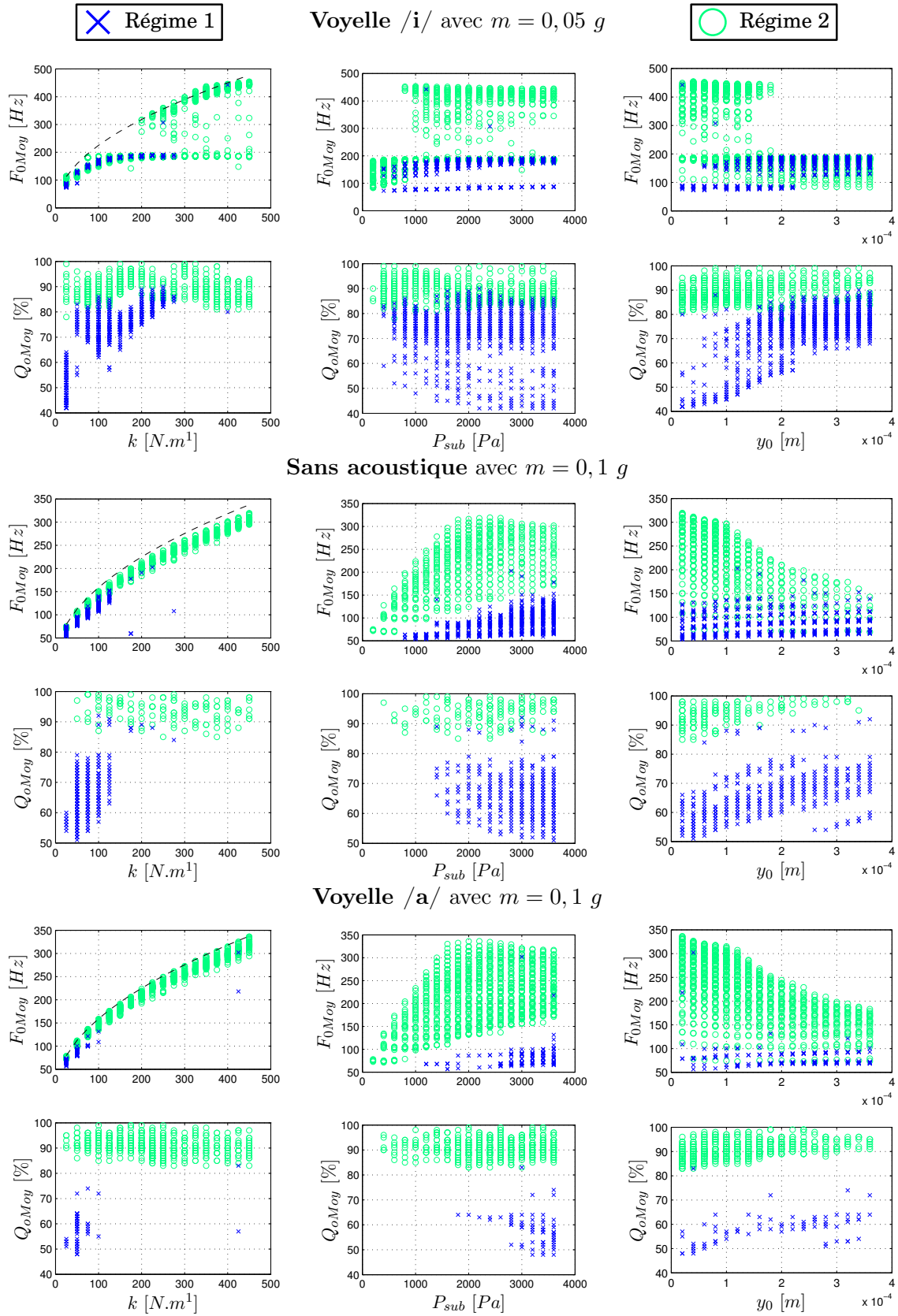


FIGURE E.4 – Fréquences fondamentales F_{0Moy} et quotients ouverts moyens Q_{oMoy} pour les régimes 1 et 2, en fonction de la raideur k , de la pression sous-glottique P_{sub} et de la demi-aperture glottique y_0 . Les courbes en pointillés représentent la fréquence de résonance mécanique « naturelle » des cordes vocales, définie par $F_m = \sqrt{k/m}/(2\pi)$.

Bibliographie

- [ABT00] F. ALIPOUR, D. A. BERRY et I. R. TITZE : A finite-element model of vocal-fold vibration. *J. Acoust. Soc. Am.*, 108(6):3003–3012, 2000.
- [AFS09] F. ALIPOUR, E. M. FINNEGAN et R. C. SCHERER : Aerodynamic and acoustic effects of abrupt frequency changes in excised larynges. *Journal of Speech, Language, and Hearing Research*, 52(2):465–481, 2009.
- [AGA13] M. ARNELA, O. GUASCH et F. ALÍAS : Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations. *The Journal of the Acoustical Society of America*, 134(4):2946–2954, 2013.
- [AKH64] A. ANGELOCCI, G. KOPP et A. HOLBROOK : The vowel formants of deaf and normal hearing eleven to fourteen year old boys. *J. Speech Hear. Dis.*, 29:156–170, 1964.
- [ALLA13] H. ARSIKERE, G. K.F. LEUNG, S. M. LULICH et A. ALWAN : Automatic estimation of the first three subglottal resonances from adults’ speech signals with application to speaker height estimation. *Speech Communication*, 55(1):51–70, 2013.
- [AS07] F. ALIPOUR et R. C. SCHERER : On pressure-frequency relations in the excised larynx. *J. Acoust. Soc. Am.*, 122(4):2296–2305, 2007.
- [AT85] F. ALIPOUR et I. R. TITZE : Viscoelastic modeling of canine vocalis muscle in relaxation. *J. Acoust. Soc. Am.*, 78(6):1939–1943, 1985.
- [AT13] M. F. ASSANEO et M. A. TREVISAN : Revisiting the two-mass model of the vocal folds. *Papers in Physics*, 5:1–7, 2013.
- [Aza92] Z. AZAMI : Inversion acoustico-articulatoire par algorithmes génétiques, 1992.
- [Bad05] R. BADEAU : *Méthodes à haute résolution pour l’estimation et le suivi de sinusoides modulées : Application aux signaux de musique*. Thèse de doctorat, ENST, 2005.
- [Bae79] T. BAER : Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes . *J. Acoust. Soc. Am.*, 65:1271, 1979.
- [Bae81] T. BAER : Observation of vocal fold vibration : Measurement of excised larynges. *Vocal fold physiology*, pages 119–133, 1981.
- [Bai09] L. BAILLY : *Interaction entre cordes vocales et bandes ventriculaires en phonation : exploration in-vivo, modélisation physique, validation in-vitro*. Thèse de doctorat, Université du Maine, 2009.
- [BAVG13] B. BERGEOT, A. ALMEIDA, C. VERGEZ et B. GAZENGEL : Prediction of the dynamic oscillation threshold in a clarinet model with a linearly increasing blowing pressure : influence of noise. *Nonlinear Dynamics*, 74(3):591–605, octobre 2013. 14 pages.

- [Ber99] D. P. BERNERS : *Acoustics and signal processing techniques for physical modeling of brass instruments*. Thèse de doctorat, stanford university, 1999.
- [BGBN87] T. BAER, J.C. GORE, S. BOYCE et P.W. NYE : Application of MRI to the analysis of speech production. *Magnetic Resonance Imaging*, 5(1):1–7, 1987.
- [BHTS96] D. A. BERRY, H. HERZEL, I. R. TITZE et B. H. STORY : Bifurcations in excised larynx experiments. *Journal of Voice*, 10(2):129–138, 1996.
- [Boo66] D. R. BOONE : Modification of the voices of deaf children. *The Volta Review*, 68:686–694, 1966.
- [Bru98] M. BRUNEAU : *Manuel d’acoustique fondamentale*. Hermès édition, 1998.
- [BS05] S. BILBAO et J. O. SMITH : Energy-conserving finite difference schemes for nonlinear strings. *Acta acustica united with acustica*, 91(2):299–311, 2005.
- [Bus81] M. A. BUSH : *Vowel articulation and laryngeal control in the speech of the deaf*. Thèse de doctorat, Massachusetts Institute of Technology, 1981.
- [Byr93] D. BYRD : 54,000 American stops. volume 83, pages 97–116. UCLA Working Papers in Phonetics, 1993.
- [Can86] J. CANNY : A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):679–698, 1986.
- [CGC00] J. S. CULLEN, J. GILBERT et D. M. CAMPBELL : Brass instruments : linear stability analysis and experiments with an artificial mouth. *Acta Acustica united with Acustica*, 86(4):704–724, 2000.
- [CH88] T. H. CRYSTAL et A. S. HOUSE : Segmental durations in connected-speech signals : current results. *Journal of Acoustical Society of America*, 83(4):1553–1573, 1988.
- [Cis08] J. CISONNI : *Modélisation et inversion d’un système complexe de production de signaux acoustiques. Application à la voix et aux pathologies*. Thèse de doctorat, Institut National Polytechnique de Grenoble-INPG, 2008.
- [CK08] A. CHAIGNE et J. KERGMARD : *Acoustique des instruments de Musique*. Belin édition, 2008.
- [CL91] D. G. CHILDERS et C.K. LEE : Vocal quality factors : Analysis, synthesis, and perception. *the Journal of the Acoustical Society of America*, 90(5):2394–2410, 1991.
- [CL99] T. CHO et P. LADEFOGED : Variation and universals in VOT : evidence from 18 languages. *Journal of phonetics*, 27(2):207–229, 1999.
- [Cra87] L. I. J. CRANEN : *The acoustic impedance of the glottis, measurements and modelling*. Thèse de doctorat, University of Nijmegen, 1987.
- [CT06] R. W. CHAN et I. R. TITZE : Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics. *J. Acoust. Soc. Am.*, 119(4):2351–2362, 2006.
- [CTT97] R. W. CHAN, I. R. TITZE et M. R. TITZE : Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa. *J. Acoust. Soc. Am.*, 101(6):3722–3727, 1997.

- [Dar87] S. N. DART : An aerodynamic study of Korean stop consonants : Measurements and modeling. *J. Acoust. Soc. Am.*, 81(1):138–147, 1987.
- [dB60] J. Van den BERG : Vocal ligaments versus registers. *Current problems in phoniatrics and Iogopedics*, 1:19–34, 1960.
- [dB68] J. Van den BERG : Register problems. *Annals of the New York Academy of Sciences*, 155(1):129, 1968.
- [dBT59] J. Van den BERG et T.S. TAN : Results of experiments with human larynxes. *ORL*, 21(6):425–450, 1959.
- [dCK02] A. de CHEVEIGNÉ et H. KAWAHARA : YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.
- [DPBL14] L. DELEBECQUE, X. PELORSON, D. BEAUTEMPS et X. LAVAL : Modélisation physique de la production de séquences voyelle-plosive-voyelle. In *12eme Congrès Français d'Acoustique*, Poitiers, France, avril 2014.
- [DPV⁺03] M. DEVERGE, X. PELORSON, C. VILAIN, P.-Y. LAGRÉE, F. CHENTOUF, J. WILLEMS et A. HIRSCHBERG : Influence of collision on the flow through in-vitro rigid models of the vocal folds. *J. Acoust. Soc. Am.*, 114(6):3354–3362, décembre 2003.
- [Duc01] E. DUCASSE : *Modélisation et simulation dans le domaine temporel d'instruments à vent à anche simple en situation de jeu : méthodes et modèles*. Thèse de doctorat, Le Mans, 2001.
- [EO79] W. G. EWAN et J. J. OHALA : Can intrinsic vowel F0 be explained by source/tract coupling? *The Journal of the Acoustical Society of America*, 66(2):358–362, 1979.
- [ESW97] J. EPPS, J. SMITH et J. WOLFE : A novel instrument to measure acoustic resonances of the vocal tract during phonation. *Measurement Science and Technology*, 8(10):1112, 1997.
- [Fab57] P. FABRE : Un procédé électrique percutanée d'inscription de l'accolement glottique au cours de la phonation : glottographie de haute fréquence, premiers résultats. *Bulletin de l'Academie Nationale de Médecine*, pages 66–69, 1957.
- [FJ72] E. FISCHER-JØRGENSEN : PTK et BDG français en position intervocalique accentuée. *Papers in linguistics and phonetics to the memory of Pierre Delattre*, pages 143–200, 1972.
- [FL68] J. L. FLANAGAN et L. L. LANDGRAF : Self-oscillating source for vocal-tract synthesizers. *Audio and Electroacoustics, IEEE Transactions on*, 16(1):57–64, 1968.
- [Fla72] J. L. FLANAGAN : *Speech analysis : Synthesis and perception*. Springer-Verlag, 1972.
- [FS11] L. P. FULCHER et R. C. SCHERER : Phonation threshold pressure : Comparison of calculations and measurements taken with physical models of the vocal fold mucosa. *J. Acoust. Soc. Am.*, 130(3):1597–1605, 2011.

- [Gar07] R. GARREL : *Contribution à la modélisation de la vibration des cordes vocales par l'application de la théorie des oscillateurs à relaxation*. Thèse de doctorat, Université de la Méditerranée, 2007.
- [GH83] J. GUCKENHEIMER et P. HOLMES : *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42. Springer Science & Business Media, 1983.
- [Gir12] M. GIROUX : Étude du rôle de la rétroaction auditive dans la production et la perception de la prosodie de l'interrogation chez des sourds postlinguistiques implantés. Mémoire de D.E.A., 2012.
- [GKP90] J GILBERT, J KERGMARD et JD POLACK : On the reflection functions associated with discontinuities in conical bores. *J. Acoust. Soc. Am.*, 87(4):1773–1780, 1990.
- [Gol78] T. GOLD : *Speech and hearing skills : A comparison between hard-of-hearing and deaf children*. Thèse de doctorat, University of New-York, 1978.
- [GPP98] J. GILBERT, S. PONTUS et J.-F. PETIOT : Artificial buzzing lips and brass instruments : Experimental results. *J. Acoust. Soc. Am.*, 104(3):1627–1632, 1998.
- [Gre56] D. S. GREEN : *Fundamental frequency of the speech of profoundly deaf individuals*. Thèse de doctorat, Purdue University, 1956.
- [Gun03] H. E. GUNTER : A mechanical model of vocal-fold collision with high spatial and temporal resolution. *The Journal of the Acoustical Society of America*, 113(2):994–1000, 2003.
- [GYR⁺06] A. GIOVANNI, P. YU, J. RÉVIS, M.D. GUARELLA, B. TESTON et M. OUAKNINE : Analyse objective des dysphonies avec l'appareillage EVA, état des lieux. *Revue Oto-Rhino-Laryngologie Française*, 90:183–192, 2006.
- [Hau03] M. HAUTEFEUILLE : Calibration statique et dynamique du système de mesures aérodynamiques EVA. Rapport technique, ENSERG - Institut National Polytechnique de Grenoble, 2003.
- [HCP09] A. Van HIRTUM, J. CISONNI et X. PELORSON : On quasi-steady laminar flow separation in the upper airways. *Communications in Numerical Methods in Engineering*, 25:447–461, 2009.
- [HCS94] M. B. HIGGINS, A. E. CARNEY et L. SCHULTE : Physiological assessment of speech and voice production of adults with hearing loss. *Journal of Speech and Hearing Research*, 37(3):510–531, jun 1994.
- [HdDC05] N. HENRICH, C. D'ALESSANDRO, B. DOVAL et M. CASTELLENGO : Glottal open quotient in singing : Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *The Journal of the Acoustical Society of America*, 117(3):1417–1430, 2005.
- [Hél02] T. HÉLIE : *Modélisation physique d'instruments de musique en systèmes dynamiques et inversion*. Thèse de doctorat, Université de Paris-Sud, 2002.
- [Hen01] N. HENRICH : *Study of the glottal source in speech and singing : Modeling and estimation, acoustic and electroglottographic measurements, perception*. Theses, Université Pierre et Marie Curie - Paris VI, novembre 2001.

- [Her14] N. HERMANT : *Observation, modélisation et simulation des vibrations des maquettes de plis vocaux. Applications à des configurations pathologiques*. Thèse de doctorat, Institut National Polytechnique de Grenoble, Oct 2014.
- [HHM03] H. HADDAR, T. HÉLIE et D. MATIGNON : A Webster-Lokshin model for waves with viscothermal losses and impedance boundary conditions : strong solutions. *In Mathematical and Numerical Aspects of Wave Propagation WAVES 2003*, pages 66–71. Springer, 2003.
- [HK95] H. HERZEL et C. KNUDSEN : Bifurcations in a vocal fold model. *Nonlinear Dynamics*, 73(1):53–64, 1995.
- [HM06a] T. HÉLIE et D. MATIGNON : Diffusive representaton for the analysis and simulation of e flared acoustic pipes with visco-thermal losses. *Mathematical Models and Methods in Applied Sciences*, 16(04):503–536, 2006.
- [HM06b] T. HÉLIE et D. MATIGNON : Representations with poles and cuts for the time-domain simulation of fractional systems and irrational transfer fonctions . *Signal Processing*, 86:2516–2528, 2006.
- [HN42] C. V. HUDGINS et F.C. NUMBERS : An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, 25:289–392, 1942.
- [Hof98] G. C. J. HOFMANS : *Vortex sound in confined flows*. Thèse de doctorat, Technische Universiteit Eindhoven, 1998.
- [Hon83] K. HONDA : Relationship between pitch control and vowel articulation. *Haskins Laboratories Status Report on Speech Research*, 73:269–282, 1983.
- [Hoo66] R. B. HOOD : *Some physical concomitants of the perception of speech rhythm of the deaf*. Thèse de doctorat, stanford university, 1966.
- [HRC⁺03] N. HENRICH, B. ROUBEAU, M. CASTELLENGO *et al.* : On the use of electroglottography for characterisation of the laryngeal mechanisms. *In Stockholm Music Acoustics Conference*. Stockholm Sweden, 2003.
- [HSW12] N. HANNA, J. SMITH et J. WOLFE : Low frequency response of the vocal tract : acoustic and mechanical resonances and their losses. *In Proc. Acoustics*, 2012.
- [HVO70] M. HIRANO, W. VENNARD et J. OHALA : Regulation of register, pitch and intensity of voice. *Folia Phoniatrica et Logopaedica*, 22(1):1–20, 1970.
- [IF72] K. ISHIZAKA et J. L. FLANAGAN : Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst. Tech. J.*, 51(6):1233–1268, 1972.
- [JSW04] E. JOLIVEAU, J. SMITH et J. WOLFE : Vocal tract resonances in singing : The soprano voice. *J. Acoust. Soc. Am.*, 116(4):2434–2439, 2004.
- [Kea84] P. A. KEATING : Aerodynamic modeling at UCLA. volume 59, pages 18–28. UCLA Working Papers in Phonetics, 1984.
- [Ken14] T. KENNY : Coughing up Blood (Haemoptysis). *Patient, Trusted medicinal information and support*, Oct 2014.
- [Ker91] J. KERGMARD : Calculation of discontinuities in waveguides using mode matching method : an alternative to the scattering matrix approach. *Journal d’Acoustique*, 4(2):111–138, 1991.

- [KFL11] L. L. KOENIG, S. FUCHS et J. C. LUCERO : Effects of consonant manner and vowel height on intraoral pressure and articulatory contact at voicing offset and onset for voiceless obstruents. *J. Acoust. Soc. Am.*, 129(5):3233–3244, 2011.
- [KL62] J. K. KELLY et C. C. LOCHBAUM : Speech synthesis. pages 1–4. 4th International Congress on Acoustics, 1962.
- [Kla75] D. H. KLATT : Voice onset time, frication, and aspiration in word-initial consonant clusters. *J. of Speech Lang Hear. Res.*, 18:686–706, 1975.
- [LA64] L. LISKER et A. S. ABRAMSON : A cross-language study of voicing in initial stops : Acoustical measurements. *Word*, 20:384–422, 1964.
- [Lam10] S. LAMESCH : *Mécanismes laryngés et voyelles en voix chantée Dynamique vocale, phonétogrammes de paramètres glottiques et spectraux, transitions de mécanismes*. Thèse de doctorat, Université Pierre et Marie Curie - Paris VI, 2010.
- [LDC57] A. M. LIBERMAN, P. C. DELATTRE et F. S. COOPER : Some cues for the distinction between voiced and voiceless stops in initial position. *J. Acoust. Soc. Am.*, 29(11):1254–1254, 1957.
- [Leh70] I. LEHISTE : Suprasegmentals. *MIT, Press*, 1970.
- [LG97] A. LÖFQVIST et V.L. GRACCO : Lip and jaw kinematics in bilabial stop consonant production. *Journal of Speech, Language, and Hearing Research*, 40(4):877–893, aug 1997.
- [LGM01] R. LAJE, T. GARDNER et G. B. MINDLIN : Continuous model for vocal fold oscillations to study the effect of feedback. *Phys. Rev. E*, 64:056201, Oct 2001.
- [LH14] N. LOPES et T. HÉLIE : A power-balanced model of a valve exciter including shocks and based on a conservative jet for brass instruments : Simulations and comparisons with standard models. *In Proceedings of the International Symposium on Musical Acoustics*, 2014.
- [LHVH98] N. J. C. LOUS, G. C. J. HOFMANS, R. N. J. VELDHUIS et A. HIRSCHBERG : A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design. *Acta Acustica*, 84:1135–1150, 1998.
- [Lie70] Ph LIEBERMAN : A study of prosodic features. *Current trends in linguistics*, 12:15–48, 1970.
- [LK05] J. C. LUCERO et L. L. KOENIG : Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control. *J. Acoust. Soc. Am.*, 117(3):1362–1372, 2005.
- [LKL⁺11] J. C. LUCERO, L. L. KOENIG, K. G. LOURENÇO, N. RUTY et X. PELORSON : A lumped mucosal wave model of the vocal folds revisited : Recent extensions and oscillation hysteresis. *J. Acoust. Soc. Am.*, 129(3):1568–1579, 2011.
- [LKM69] P. LIEBERMAN, R. KNUDSON et J. MEAD : Determination of the Rate of Change of Fundamental Frequency with Respect to Subglottal Air Pressure During Sustained Phonation. *J. Acoust. Soc. Am.*, 45(6):1537–1543, 1969.
- [LLDE75] L. LISKER, A.M. LIBERMAN, D. DECHOWITZ et D.M. ERICKSON : On pushing the voice-onset-time (VOT) boundary about. *J. Acoust. Soc. Am.*, 57(S1):S50–S50, 1975.

- [LLH⁺12] J. C. LUCERO, K. G. LOURENÇO, N. HERMANT, A. Van HIRTUM et X. PELORSON : Effect of source–tract acoustical coupling on the oscillation onset of the vocal folds. *J. Acoust. Soc. Am.*, 132(1):403–411, 2012.
- [LM05] R. LAJE et G. B. MINDLIN : Modeling source-source and source-filter acoustic interaction in birdsong. *Physical Review E*, 72(3):036218, 2005.
- [Luc93] J. C. LUCERO : Dynamics of the two-mass model of the vocal folds : Equilibria, bifurcations, and oscillation region. *J. Acoust. Soc. Am.*, 94(6):3104–3111, 1993.
- [Luc96] J. C. LUCERO : Chest- and falsetto-like oscillations in a two-mass model of the vocal folds. *J. Acoust. Soc. Am.*, 100(5):3355–3359, 1996.
- [Luc98] J. C. LUCERO : Optimal glottal configuration for ease of phonation. *Journal of Voice*, 12(2):151–158, 1998.
- [Luc99] J. C. LUCERO : A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset–offset . *J. Acoust. Soc. Am.*, 105(1), 1999.
- [Luc05] J. C. LUCERO : Bifurcations and limit cycles in a model for a vocal fold oscillator. *Communications in Mathematical Sciences*, 3(4):517–529, 12 2005.
- [Lul53] Hans LULLIES : Physiologie der stimme und sprache. In *Lehrbuch der Physiologie*, pages 163–273. Springer, 1953.
- [LW91] H. LANE et J. W. WEBSTER : Speech deterioration in postlingually deafened adults. *The Journal of the Acoustical Society of America*, 89(2):859–866, 1991.
- [MA88] J MARTINEZ et J AGULLÓ : Conical bores. Part I : Reflection functions associated with discontinuities. *J. Acoust. Soc. Am.*, 84(5):1613–1619, 1988.
- [Mai11] B. MAILLOU : Approche physique de la voix d’une personne sourde. Mémoire de D.E.A., Université du Maine, juillet 2011.
- [Mar68] J. MARTONY : On correction of voice pitch level for severely hard of hearing subjects. *Am. Ann. Deaf*, 113:195–202, 1968.
- [Mat94] D. MATIGNON : *Représentations en variables d’état de modèles de guides d’ondes avec dérivation fractionnaire*. Thèse de doctorat, 1994.
- [Meu07] C. MEUNIER : Phonétique acoustique. In AUZOU P., éditeur : *Les dysarthries*, pages 164–173. Solal, 2007.
- [MEV79] R. B. MONSEN, A. M. ENGBRETSON et N. VEMULA : Some effects of the deafness on the generation of voice. *J. Acoust. Soc. Am.*, 66:1680–1690, 1979.
- [Mey97] E. A. MEYER : Zur Tonbewegung des Vokals im gesprochenen und gesungenen einzelwort. *Phonetische Studien (Beiblatt zu der Zeitschrift Die Neuren Sprachen)*, 10:1–21, 1897.
- [MG83] N. S. MCGARR et C. E. GELFER : Simultaneous measurements of vowels produced by a hearing-impaired speaker. *Language and speech*, 26(3):233–246, 1983.
- [MH80] N. S. MCGARR et K. S. HARRIS : Articulatory control in a deaf speaker. *Haskins Laboratories Status Report on Speech Research*, pages 307–330, 1980.
- [MH97] P. MERGELL et H. HERZEL : Modelling biphonation—the role of the vocal tract. *Speech Communication*, 22(2):141–154, 1997.

- [MHM08] R. MIGNOT, T. HÉLIE et D. MATIGNON : Stable realization of a delay system modeling a convergent acoustic cone. pages 1574–1579. Mediteranean conference on control and automation, 2008.
- [Mig09] R. MIGNOT : *Réalisation en guides d'ondes numériques stables d'un modèle acoustique réaliste pour la simulation en temps-réel d'instruments à vent*. Thèse de doctorat, Telecom ParisTech, 2009.
- [Mil00] D. G. MILLER : *Registers in singing : empirical and systematic studies in the theory of the singing voice*. Thèse de doctorat, University of Groningen, 2000.
- [MKL95] R. S. MCGOWAN, L. L. KOENIG et A. LÖFQVIST : Vocal tract aerodynamics in /aCa/ utterances : Simulations. *Speech Communication*, 16:67–68, 1995.
- [Mon78] R. B. MONSEN : Toward measuring how well hearing-impaired children speak. *Journal of Speech, Language, and Hearing Research*, 21(2):197–219, 1978.
- [MW80] E. M. MÜLLER et Jr W. S. BROWN : Variations in the supraglottal air pressure waveform and their articulatory interpretation. *Speech and Language : Advances in Basic Research and Practice*, 4:317–389, 1980.
- [Ned98] C. J. NEDERVEEN : *Acoustical aspects of woodwind instruments*. Northern Illinois University Press, 1998.
- [Nic75] R. S. NICKERSON : Characteristics of the Speech of Deaf Persons. *Volta Review*, 77(6):342–62, 1975.
- [Nob67] E. N. NOBER : Articulation of the deaf. *Exceptional children*, 33:611–621, 1967.
- [NS89] A.N. NORRIS et I.C. SHENG : Acoustic radiation from a circular pipe with an infinite flange. *Journal of Sound and Vibration*, 135(1):85–93, 1989.
- [OE76] J. J. OHALA et B. W. EUKEL : Explaining the intrinsic pitch of vowels. *J. Acoust. Soc. Am.*, 60(S1):S44–S44, 1976.
- [OLS79] M. J. OSBERGER, H. LEVITT et R. SLOSBERG : Acoustic characteristics of correctly produced vowels in deaf children's speech. *J. Acoust. Soc. Am.*, 66(S13), 1979.
- [OM82] M. J. OSBERGER et N. S. MCGARR : Speech production characteristic of the hearing impaired. *Speech and Language : Advances in Basic Research and Practice*, 8:221–283, 1982.
- [Osb81] M. J. OSBERGER : Fundamental frequency characteristics of the speech of the hearing impaired. *J. Acoust. Soc. Am.*, 69(S1):S68–S68, 1981.
- [Per69] J. S. PERKELL : *Physiology of speech production : results and implications of a quantitative cineradiographic study*. The mit press édition, 1969.
- [PG99] J-F PETIOT et J. GILBERT : Aided design of brass musical instruments. *In Integrated Design and Manufacturing in Mechanical Engineering'98*, pages 497–504. Springer, 1999.
- [PHRB97] X. PELORSON, G.C.J. HOFMANS, M. RANUCCI et R.C.M. BOSCH : On the fluid mechanics of bilabial plosives. *Speech Communication*, 22:155–172, 1997.

- [PHvH⁺94] X. PELORSON, A. HIRSCHBERG, R. R. van HASSEL, A. P. J. WIJNANDS et Y. AUREGAN : Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model. *J. Acoust. Soc. Am.*, 96(6):3416–3431, 1994.
- [PL12] X. PELORSON et X. LAVAL : An experimental replica of the vocal folds to study normal and pathological voice. In Société Française D’ACOUSTIQUE, éditeur : *Proceedings of the Acoustics 2012 Nantes Conference*, Nantes, France, avril 2012. Proceedings of the Acoustics 2012 Nantes Conference.
- [PML⁺97] J. PERKELL, M. MATTHIES, H. LANE, F. GUENTHER, R. WILHELMS-TRICARICO, J. WOZNIAC et P. GUIOD : Speech motor control : Acoustic goals, saturation effects, auditory feedback and internal models. *Speech communication*, 22(2):227–250, 1997.
- [Pol91] J. D. POLACK : Time domain solution of Kirchhoff’s equation for sound propagation in viscothermal gases : a diffusion process. *J. Acoust. Soc. Am.*, 4:47–67, 1991.
- [PVC⁺96] X. PELORSON, C. VESCOVI, E. CASTELLI, A. HIRSCHBERG, A. P. J. WIJNANDS et H. M. A. BAILLIET : Description of the Flow Through In-Vitro Models of the Glottis During Phonation. Application to Voiced Sounds Synthesis. *Acta Acustica united with Acustica*, 82(2):358–361, 1996.
- [RCMAG87] B ROUBEAU, C CHEVRIE-MULLER et C ARABIA-GUIDET : Electroglottographic study of the changes of voice registers. *Folia Phoniatrica et Logopaedica*, 39(6): 280–289, 1987.
- [RHC09] B. ROUBEAU, N. HENRICH et M. CASTELLENGO : Laryngeal vibratory mechanisms : The notion of vocal register revisited. *Journal of Voice*, 23(4):425–438, 2009.
- [RK89] R. ROY et T. KAILATH : ESPRIT-estimation of signal parameters via rotational invariance techniques. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(7):984–995, 1989.
- [RLG03] J. RYALLS, A. LAROCHE et F. GIROUX : Acoustic comparison of CV syllables in French-speaking children with normal hearing, moderate-to-severe and profound hearing impairment. *Journal of Multilingual Communication Disorders*, 1(2):99–114, 2003.
- [Rot68] M. ROTHENBERG : *The breath-stream dynamics of simple-released-plosive production*. Karger, 6 édition, 1968.
- [Rou93] B. ROUBEAU : *Mécanismes vibratoires laryngés et contrôle neuro-musculaire de la fréquence fondamentale*. Thèse de doctorat, Université Paris XI, 1993.
- [RPH⁺07] N. RUTY, X. PELORSON, A. Van HIRTUM, I. LOPEZ-ARTEAGA et A. HIRSCHBERG : An in vitro setup to test the relevance and the accuracy of low-order vocal folds models. *J. Acoust. Soc. Am.*, 121(1):479–490, 2007.
- [RPK86] R. ROY, A. PAULRAJ et T. KAILATH : ESPRIT–A subspace rotation approach to estimation of parameters of cisoids in noise. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(5):1340–1342, 1986.

- [RTZJ08] M. F. REGNER, C. TAO, P. ZHUANG et J. J. JIANG : Onset and offset phonation threshold flow in excised canine larynges. *Laryngoscope*, 118:1313–1317, 2008.
- [Rut07] N. RUTY : *Modèles d'interactions fluide/parois dans le conduit vocal. Applications aux voix et aux pathologies*. Thèse de doctorat, Institut National Polytechnique de Grenoble, Novembre 2007.
- [SA15] D. SCIAMARELLA et G. ARTANA : Relaxation to one-dimensional postglottal flow in a vocal fold model. *Speech Communication*, 66:176–181, 2015.
- [SD04] D. SCIAMARELLA et C. D'ALESSANDRO : On the acoustic sensitivity of a symmetrical two-mass model of the vocal folds to the variation of control parameters. *Acta Acustica*, 90:746–761, 2004.
- [SG99] H. SCHLICHTING et K. GERSTEN : *Boundary Layer theory*. Springer Verlag, 6 édition, 1999.
- [SGK⁺09] F. SILVA, P. GUILLEMAIN, J. KERGOMARD, B. MALLARONI et A. N. NORRIS : Approximation formulae for the acoustic radiation impedance of a cylindrical pipe. *Journal of Sound and Vibration*, 322:255–263, 2009.
- [SH95] I. STEINECKE et H. HERZEL : Bifurcations in an asymmetric vocal-fold model. *J. Acoust. Soc. Am.*, 97(3):1875–1884, 1995.
- [Sir13] A. SIRDEY : *Analyse modale de sons d'impact par méthodes haute résolution pour la catégorisation perceptive des matériaux*. Thèse de doctorat, 2013.
- [SK74] K. N. STEVENS et D. H. KLATT : Role of formant transitions in voiced-voiceless distinction for stops. *J. Acoust. Soc. Am.*, 55(3):653–659, 1974.
- [Smi75] C. R. SMITH : Residual hearing and speech production in deaf children. *Journal of Speech, Language, and Hearing Research*, 18(4):795–811, 1975.
- [Smi96] J. O. SMITH : Physical modeling synthesis update. *Computer Music Journal*, pages 44–56, 1996.
- [SNBR76] K. N. STEVENS, R. S. NICKERSON, A. BOOTHROYD et A. M. ROLLINS : Assessment of nasalization in the speech of deaf children. *Journal of Speech, Language, and Hearing Research*, 19(2):393–416, 1976.
- [SNR78] K. N. STEVENS, R. S. NICKERSON et A. M. ROLLINS : On describing the suprasegmental properties of the speech of deaf children. *Advance in Prosthetic devices for the deaf : a technical report*, pages 134–155, 1978.
- [SP94] J. SVEC et J. PESÁZK : Vocal Breaks from the Modal t0 Falsetto Register. *Folia Phoniatiica et Logopaedica*, 46(2):97–103, 1994.
- [SSM99] J. G. SVEC, H. K. SCHUTTE et D. G. MILLER : On pitch jumps between chest and falsetto registers in voice : Data from living and excised human larynges. *The Journal of the Acoustical Society of America*, 106(3):1523–1531, 1999.
- [Ste93] K. N. STEVENS : Models for the production and acoustics of stop consonants. *Speech communication*, 13(3):367–375, 1993.
- [Sto95] B. H. STORY : *Physiologically-Based Speech Simulation Using AN Enhanced Wave-Reflection Model of the Vocal Tract*. Thèse de doctorat, University of Iowa, 1995.

- [TG95] B. TESTON et B. GALINDO : A diagnostic and rehabilitation aid workstation for speech and voice pathologies. pages 18–21. Fourth European Conference on Speech Communication and Technology, ISCA, September 1995.
- [TGG99] B. TESTON, A. GHIO et B. GALINDO : A multisensor data acquisition and processing system for speech production investigation. *In Proceedings of International Congress of Phonetic Sciences (ICPhS)*, pages 2251–2254, 1999.
- [THvH07] I. T. TOKUDA, J. HORÁČEK, J. G. ŠVEC et H. HERZEL : Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments. *The Journal of the Acoustical Society of America*, 122(1):519–531, 2007.
- [Tit88a] I. R. TITZE : A framework for the study of vocal registers. *Journal of Voice*, 2(3):183–194, 1988.
- [Tit88b] I. R. TITZE : The physics of small-amplitude oscillation of the vocal folds . *J. Acoust. Soc. Am.*, 83(4):1536–1552, 1988.
- [Tit89] I. R. TITZE : On the relation between subglottal pressure and fundamental frequency in phonation. *J. Acoust. Soc. Am.*, 85(2):901–906, 1989.
- [Tit94] I. R. TITZE : *Principles of Voice Production*. Prentice Hall, 1994.
- [TJZ06] C. TAO, J. J. JIANG et Y. ZHANG : Simulation of vocal fold impact pressures with a self-oscillating finite-element model. *J. Acoust. Soc. Am.*, 119(6):3987–3994, 2006.
- [TST95] I. R. TITZE, S. S. SCHMIDT et M. R. TITZE : Phonation pressure threshold pressure in a physical model of the vocal fold mucosa. *J. Acoust. Soc. Am.*, 97(5):3080–3084, 1995.
- [TZKH10] I. T. TOKUDA, M. ZEMKE, M. KOB et H. HERZEL : Biomechanical modeling of register transitions and the role of vocal tract resonators. *J. Acoust. Soc. Am.*, 127(3):1528–1536, 2010.
- [Vai06] J. VAISSIÈRE : *La phonétique, Que-Sais-Je ?* Presses Universitaires de France, 2006.
- [Val95a] V. VALIMAKI : *Discrete-time modeling of acoustic tubes using fractional delay filters*. Thèse de doctorat, Helsinki University of Technology, 1995.
- [VAL95b] E. VILKMAN, P. ALKU et A-M. LAUKKANEN : Vocal-fold collision mass as a differentiator between registers in the low-pitch range. *Journal of Voice*, 9(1):66–73, 1995.
- [Ven67] W. VENNARD : *Singing : the mechanism and the technic*. Carl Fischer, LLC, 1967.
- [Vil02] C. VILAIN : *Contribution à la synthèse de parole par modèle physique. Application à l'étude des voix pathologiques*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 2002.
- [Voe35] C. H. VOELKER : A preliminary strobophotoscopic study of the deaf. *Am. Ann. Deaf*, 80:243–259, 1935.

- [Wal90] R. S. WALDSTEIN : Effects of postlingual deafness on speech production : implications for the role of auditory feedback. *J. Acoust. Soc. Am.*, 88(5):2099–2114, 1990.
- [Web19] A. G. WEBSTER : Acoustical impedance and the theory of horns and of the phonograph. *Proc Natl Acad Sci USA*, 5(7):275–282, 1919.
- [Wes83] J. R. WESTBURY : Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *J. Acoust. Soc. Am.*, 73(4):1322–1336, 1983.
- [WK86] J. R. WESTBURY et P.A. KEATING : On the naturalness of stop consonant. *Journal of Linguistics*, 22:145–166, 1986.
- [WL95] D. H. WHALEN et A. G. LEVITT : The universality of intrinsic F0 of vowels. *Journal of Phonetics*, 23(3):349–366, 1995.
- [Yao09] Y. YAO : Understanding VOT variation in spontaneous speech. In *Proceedings of the 18th International Congress of Linguists (CIL XVIII)*, 2009.
- [ZNB09] Z. ZHANG, J. NEUBAUER et D. A. BERRY : Influence of vocal fold stiffness and acoustic loading on flow-induced vibration of a single-layer vocal fold model. *Journal of sound and vibration*, 322(1):299–313, 2009.
- [Zor73] W. E. ZORUMSKI : Generalized radiation impedances and reflection coefficients of circular and annular ducts. *J. Acoust. Soc. Am.*, 54(6):1667–1673, 1973.
- [Zue76] V. ZUE : *Acoustic characteristics of stop consonants : A controlled study*. Thèse de doctorat, Massachusetts Institute of Technology, 1976.

Résumé

L'apprentissage du langage parlé nécessite un contrôle musculaire très précis des différents organes intervenant dans la production de la parole. La production de sons voisés, qui résulte de l'auto-oscillation des cordes vocales, est notamment influencée par l'ensemble du système phonatoire, du diaphragme jusqu'aux lèvres. Les travaux réalisés dans le cadre de cette thèse s'inscrivent dans un contexte de modélisation physique de la parole. Les objectifs s'articulent autour de la compréhension des phénomènes physiques gouvernant la production de sons voisés. Les études sont appliquées à des cas pour lesquels le contrôle de la production est fortement altéré, lorsque le locuteur souffre de pertes auditives importantes. Dans ce cas de figure, les interactions physiques peuvent jouer un rôle important dans l'apparition de troubles de la production. La démarche adoptée consiste alors dans un premier temps à observer les phénomènes étudiés au moyen de mesures *in vivo* puis à proposer des modèles théoriques mécanique, aérodynamique et acoustique permettant de les décrire. Ensuite, les modèles sont validés en comparant avec des mesures réalisées sur une maquette de l'appareil phonatoire. Finalement, des simulations numériques temporelles basées sur un modèle à deux masses pour décrire le comportement mécanique des cordes vocales, permettent de tester les modèles physiques pour des cas concrets de production.

La première étude se concentre sur les sauts de fréquence fondamentale qui accompagnent les transitions involontaires entre deux mécanismes laryngés lors de la production d'une voyelle. Les travaux expérimentaux et numériques montrent qu'une transition de mécanisme laryngé est la manifestation d'une bifurcation du système laryngé et que ces bifurcations se produisent lors d'une variation de la raideur des cordes vocales, de la pression sous-glottique, de l'aire glottique initiale ou bien de la longueur des résonateurs acoustiques. Les modèles théoriques permettent de reproduire les sauts de fréquence fondamentale observés expérimentalement. Ils sont utilisés pour étudier les différentes stratégies motrices responsables de ces sauts de fréquence. La deuxième étude porte sur la production de consonnes plosives, en particulier sur les effets de la réalisation d'une occlusion du conduit vocal sur l'arrêt puis l'apparition de l'oscillation des cordes vocales. Les simulations de séquences voyelle - plosive bilabiale non voisée (/p/)-voyelle effectuées montrent que l'expansion passive de la cavité supraglottique est à l'origine du maintien de l'auto-oscillation des cordes vocales après la fermeture du conduit vocal et que l'augmentation de la longueur du conduit vocal a pour effet de réduire le délai entre le relâchement de l'occlusion et l'apparition de l'oscillation des cordes vocales. Ces résultats impliquent que l'articulation joue un rôle considérable sur le mode de voisement (voisée ou non voisée) de la consonne et sur la valeur du *Voice Onset Time* pour une consonne plosive non voisée.

Abstract

Language learning requires specific muscle control of all organs that contribute to speech production. Voiced sounds production, which results from vocal folds self oscillation, is especially influenced by the whole phonatory apparatus, from diaphragm to lips. The general background of this thesis is the physical modeling of speech production and the objectives are motivated by a better comprehension of physical phenomena occurring in voiced sounds production. In the frame of this work, studies are focused on cases where speech production control is impaired, for example when the speaker suffer from an important hearing loss. In this situation, physical interactions can play an important role in speech production disorders emergence. The approach adopted here is first to observe the studied phenomena thanks to *in vivo* measurements then to describe them thanks to theoretical models. Thereafter, the models are validated by comparing theoretical results with measurements performed on a replica of the phonatory apparatus. Finally, numerical simulations in the time domain, based on a two-mass model, allow to apply physical models to specific speech production occurrences. The first study deals with fundamental frequency jumps that are observed during an involuntary transition between two different laryngeal mechanisms in case of vowel production. Experimental and numerical results highlight that a transition between two different laryngeal mechanisms is a symptom of the laryngeal system bifurcation, and that such bifurcation occurs during a variation of the vocal folds stiffness, the subglottal pressure, the prephonatory glottal area or the acoustic resonators length. The theoretical models allow to simulate the fundamental frequency jumps that are observed experimentally. They are used to study the different motor strategies responsible for these frequency jumps. The second study deals with plosive consonants production, and in particular with the effects of a vocal tract occlusion on voicing offset and onset. Simulations of vowel – voiceless plosive - vowel production highlight that passive expansion of the supraglottal cavity is responsible for the voicing extension after vocal tract closure, and that increase of the vocal tract length leads to a shorter delay between the vocal tract occlusion release and the voicing onset. These results highlight that the articulation plays an important role in voicing (voiced or voiceless) and in voice-onset-time value for a voiceless plosive.
